TECHNOMETRICS
1
5
19
27
43
55
67
83
103
109
115
117
121

TECHNOMETRICS

A Journal of Statistics for the Physical, Chemical and Engineering Sciences

Engin. Library
QA
276
.A1
T25

v. 2
no. 2

MAY 1960

VOLUME 2 NUMBER 2

Published quarterly by:

The American Society for Quality Control and The American Statistical Association

THE UNIVERSITY OF MICHIGAN LIBRARIES

TECHNOMETRICS

The purpose of **Technometrics** is to contribute to the development and use of statistical methods in the physical, chemical and engineering sciences. This objective places a high premium on succinct communication among the physical scientist, engineer, statistician and mathematician. The journal will accept for publication papers describing new statistical techniques expected to be useful in these sciences, papers illustrating the application of known statistical methods to new or novel environments, expository or tutorial papers on particular statistical methods, and papers dealing with the philosophy and problems of applying statistical methods to research, development, design and performance. Brief descriptions of problems requiring solution and short technical notes will also be accepted for publication. Letters to the Editor, signed by the author and limited in length will be published when they are considered timely and appropriate. All papers should contain a short but clear summary of contents and conclusions, an expository section containing numerical examples whenever practicable, and appropriate additional sections relating to technical derivations.

Subscription Rates

The annual subscription rate for members of the American Society for Quality Control or the American Statistical Association is $6.00 a year. The annual subscription rate for non-members is $8.00 a year.

Members of the sponsoring societies, the American Society for Quality Control and the American Statistical Association, may subscribe to the journal while paying their annual dues or by check or money order made out to Technometrics and mailed to:

TECHNOMETRICS
Post Office Box 587
Benjamin Franklin Station
Washington 4, D. C.

All non-member subscriptions should be mailed to this address. Communications concerning changes of address, subscriptions, back numbers, etc., should also be sent to this address. Whenever possible, a copy of the address taken from an issue of the periodical should accompany a change of address request. All subscription fees are payable in the currency of the United States of America.

Communication concerning *membership* in either of the sponsoring societies should be sent to that society.

Second class postage paid at Richmond, Va.

TECHNOMETRICS

A Journal of Statistics for the Physical, Chemical and Engineering Sciences

Published Quarterly by

THE AMERICAN SOCIETY FOR QUALITY CONTROL

and the

AMERICAN STATISTICAL ASSOCIATION

Editor

J. STUART HUNTER

Associate Editors

WILLIAM ALLEN
G. A. BARNARD
C. A. BENNETT
CUTHBERT DANIEL
BESSE B. DAY
R. J. HADER
MARTIN WILK
MARVIN ZELEN

Management Committee

PAUL S. OLMSTEAD, *Chairman*

For the American Society for Quality Control

J. Y. MCCLURE
MAYNARD RENNER
H. L. WEHRLY

For the American Statistical Association

IRVING BURR
ALMARIN PHILLIPS
DONALD C. RILEY

Published Quarterly in February, May, August and November by the Technometrics Management Committee of the American Society for Quality Control and the American Statistical Association. Editorial Office: Mathematics Research Center, U. S. Army; The University of Wisconsin, Madison, Wisconsin. Publication Office: Wm. Byrd Press, P. O. Box 2W, Richmond 5, Virginia. Second class mailing privileges granted at Richmond, Virginia.

Composed and Printed at the

WILLIAM BYRD PRESS, INC., RICHMOND, VIRGINIA, U.S.A.

CONTENTS

TECHNOMETRICS, VOL. 2, NO. 2, MAY 1960

Rejection of Outliers ..... *F. J. Anscombe* 123

Locating Outliers in Factorial Experiments..... *C. Daniel* 149

Discussion of the Papers of Messrs. Anscombe and Daniel
W. H. Kruskal, T. S. Ferguson, J. W. Tukey and E. J. Gumbel 157

Tests for the Validity of the Assumptions that the Underlying Distribution of Life is Exponential, Part II ..... *B. Epstein* 167

Partial Duplication of Response Surface Designs.... *O. Dykstra* 185

A Rank Sum Test for Comparing all Pairs of Treatments
R. G. D. Steel 197

The Percentile Points of Distributions Having Known Cumulants ..... *R. A. Fisher and E. A. Cornish* 209

An Approximation to the Negative Moments of the Positive Binomial Useful in Life Testing
W. Mendenhall and E. H. Lehman, Jr. 227

Order Statistics from the Gamma Distribution .... *S. S. Gupta* 243

Parallel Fractional Replicates ..... *C. Daniel* 263

Notices 269

Rejection of Outliers*

F. J. ANSCOMBE

Princeton University and University of Chicago

with collaboration by Irwin Guttman

McGill University

If one reading is a long way from the rest in a series of replicate determinations, or if in a least-squares analysis one reading is found to have a much greater residual than the others, there is temptation to reject it as spurious. Numerous criteria for the rejection of outliers have been proposed and discussed during the past 100 years. They seem always to have been regarded as something like significance tests, and attention has been focussed on rejection rates. It is suggested that rejection rules are not significance tests but insurance policies, and attention would be better focussed on error variance. A detailed study is made of the effect of routine application of rejection criteria to replicate determinations of a single value. Determinations in triplicate and quadruplicate are especially considered. Complex patterns of observations are also considered, especially factorial arrangements with high symmetry, and there is a study of the correlations between residuals. Attention is focussed mainly on rejection rules appropriate when the population variance is known, but some consideration is also given to Studentized rules.

1. SPURIOUS OBSERVATIONS

Variability or dispersion in a set of observations can arise from several different sources. Suppose that, for some reason we need not go into, it is desired to investigate the height (stature) of persons employed at a particular place. Three sources of variability in the readings are:

(i) *Inherent variability*, the variability of stature that would be observed in the population even if all measurements were perfectly accurate. This variability cannot be reduced without changing the population itself, the object of study. If we are interested in the *mean* stature of the population, we may refer to the variability as "error," since it gives rise to estimation error. But the name is misleading.

(ii) *Measurement error*, the error in using the measuring instruments. If all readings are made to the nearest centimeter, the measurement error should not exceed half a centimeter, but in fact it sometimes does. One may count as a measurement error any arithmetical mistake in reducing the original notebook entries to the form in which they are quoted as observations.

* Research carried out in part at Princeton University and in part at the Department of Statistics, University of Chicago, under sponsorship of the Logistics and Mathematical Statistics Branch, Office of Naval Research. Reproduction in whole or in part is permitted for any purpose of the United States Government.

(iii) *Execution error*, a name we may use to include any discrepancy between what we intend to do and what is actually done, other than error in the use of measuring instruments. To include in the sample of measurements the height of some person not belonging to the population, to measure something other than height, or to select a biased sample—these would be classed as errors of execution.

No observations are absolutely trustworthy. In no field of observation can we entirely rule out the possibility that an observation is vitiated by a large measurement or execution error. If a reading is found to lie a very long way from its fellows in a series of replicate observations, there must be a suspicion that the deviation is caused by a blunder or gross error of some kind. Several possible reasons why a reading might be grossly wrong can usually be thought of without difficulty. In such cases, the reading will be checked or repeated if that is possible. If not, it may be rejected as spurious because of its big residual, even though there is no other known reason for suspecting it. In sufficiently extreme cases, no one hesitates about such rejections. The question is, where should the line be drawn? One sufficiently erroneous reading can wreck the whole of a statistical analysis, however many observations there are.

Statements to the above effect have appeared many times in the literature. Even writers who have expressed total disapproval of the rejection of outliers may be found to insert a parenthetical remark, "except for obviously incorrect readings." Reasonable argument can occur, not over the permissibility of ever rejecting outliers, but over the completeness of the rejection, how far the fact of a rejection can be forgotten—as well as over the question of where precisely to draw the line. If we could be sure that an outlier was caused by a large measurement or execution error which could not be rectified (and if we had no interest in studying such errors for their own sake), we should be justified in entirely discarding the observation and all memory of it. The act of observation would have failed; there would be nothing to report. Such an observation could justly be described as *spurious*. If, on the other hand, we could be sure that an outlier was caused, not by any large error, but by some peculiarity (nonnormality) of the inherent variability of the population under study, then it might still make good sense to discard the observation from a statistical analysis based on the method of least squares, but the observation should not be forgotten. A correct statistical summary of the observations would include a report about outliers as well as an analysis of the remaining observations (cf. Kruskal, 1960).

If a rather pronounced degree of nonnormality is expected, it may be advisable to transform the observations, or to work with medians instead of means, or some other such device.* A modified least-squares method, with weights depend-

* The suggestion about medians in this context goes back to Edgeworth (1887). It is not always satisfactory in principle to make use of any such device. If the subject of study is the output of an agricultural or industrial process, where most of the variability is inherent, the mean yield is inescapably what ought to be measured, not the median yield or mean logarithmic yield (for instance). Perhaps it should be explicitly stated that in this paper we are concerned only with observations whose primary purpose is the estimation of location parameters of some sort. Observations made primarily for the measurement of dispersion are not considered.

ing on the residuals, has been suggested by Jeffreys, with particular reference to seismology and astronomy. Although such a technique is possible as a guard against spurious observations, it will not be considered here. *We envisage a situation where the intrinsic variability and ordinary measurement errors result in a nearly enough normal pattern of variation for simple averages or equal-weighted least-squares estimates to be attractive, but where it is feared that occasionally an observation will be affected by a gross error.*

An observation with an abnormally large residual will be referred to as an *outlier*. Other terms in English are "wild," "straggler", "sport" and "maverick"; one may also speak of a "discordant," "anomalous" or "aberrant" observation. *Spurious* will mean: affected by an abnormally large measurement or execution error. Usually it is not possible to check directly whether an observation is spurious; it is a matter of conjecture.

The object of this paper is to study the effect of unthinking routine use of a specified rejection rule on the estimation of means (or linear combinations of means). There are circumstances where the problem of spurious observations can appropriately be met by an impartial rejection procedure, which is thought of as part of the experimental technique. It is not suggested that the problem should always be met this way, but the results of the present study should be helpful (though not decisive) to anyone who wishes to judge individual outliers "on their merits," intelligently.

The gist of the paper can be had by reading as far as Section 3, skimming Sections 4 and 5, and then turning to Section 10. Section 7 on complex designs is to some extent a digression, and may have interest apart from the subject of outliers.

2. Sketch of the History of Rejection of Outliers

The subject of rejection rules has been held to be important and interesting in many sciences—astronomy, geodesy, chemistry, physics, ballistics. Nowadays it may be thought of as lying within the broader subject of "data processing," an essential though perhaps too little considered branch of statistical analysis. It has been discussed in countless books on the combination of observations by least squares, over a period of nearly a hundred years, as well as in many contemporary books on statistics.

One of the first references to the rejection of outliers seems to have been a remark by the leader of the German school of astronomers, Bessel, in a geodetic work published in 1838, to the effect that he had never rejected an observation merely because of its large residual; all completed observations, with equal weight, ought to be allowed to contribute to the result. "We have believed that only through strict observance of this rule could we remove arbitrariness from our results."

The first attempt at a rejection criterion based on some sort of probability reasoning was that of Peirce (1852). Peirce's argument was reproduced by Chauvenet (1863), who then gave a similar rule based on a simpler argument. A heated discussion was provoked by these rules, and by some other suggestions put forward for disposing of outliers. In due course the topic became standard

in books dealing with least squares; lengthy historical surveys were given by Czuber (1891) and Wellisch (1909). See also Rider (1933).

The next significant step after Peirce and Chauvenet was apparently made by Wright (1884). After an excellent general discussion of the problems raised by the occurrence of outliers among astronomical readings, he suggests that the best rule for a computer to follow who is not the observer is to reject any observation whose residual exceeds in magnitude five times the probable error (i e. 3.37 times the standard deviation). The reason given for this is that if the Gaussian law of error is truly satisfied, only about one observation in a thousand will be rejected, "and therefore little damage will be done in any case."

From 1925 onwards statisticians have paid much attention to the subject. Student (1927) gave an interesting account of analytic chemical determinations, and proposed the use of a range criterion. Thompson (1935) may be said to have "Studentized" Wright's rule. A good survey of this literature has been given by Grubbs (1950). A recent development is the Bliss-Cochran-Tukey rule (1956), developed for use with several of the bioassays in the U. S. Pharmacopeia.

Rejection of outliers seems to have been a peculiarly American topic. At any rate, Peirce, Chauvenet and Wright were Americans, and most of the recent developments have occurred in the U. S. The subject arose in the context of least squares, or, as current jargon has it, Model I analysis of variance; and it is curious that (apparently) it was not considered by Gauss.

Chauvenet, after he had reproduced Peirce's argument, introduced his own simplified version of the rule by some remarks beginning: "The above investigation of the criterion involves some principles, derived from the theory of probabilities, which may seem obscure to those not familiar with that branch of science." The same may be said of Chauvenet and of everyone else; and familiarity with that branch of science does not remove the obscurity. It is easy now to laugh at Peirce and Chauvenet, and at other eminent nineteenth century figures such as Glaisher and Bertrand, as they wallow in the probability quagmire. But have today's statisticians been so much clearer headed? All published proposals for rejection criteria, based on any kind of mathematical reasoning, from Peirce's onwards, have an unexplained starting point or objective, presented as though it were the only obvious one and in fact utterly obscure.

All writers seem to have thought of the problem as something in the nature of significance testing.* They have varied in the significance level judged appropriate for action, but all have regarded significance levels (or something of the sort) as relevant beyond discussion. Peirce and Chauvenet thought in terms of something like a 50% significance level per set of data. If applied to a simple sample of n observations drawn from a homogeneous normal source, either criterion rejects as often as not, very roughly, so the rejection rate is

* Jeffreys puts this the other way round and suggests that the general use of tail-area probabilities by statisticians for significance testing goes back to Chauvenet's criterion (Jeffreys, 1939, p. 316; 1948, p. 357).

something like 1 per $2n$ observations. Wright's rule rejects at something like the rate of 1 per 1000 observations, almost independently of n. Most modern statisticians, bemused by 5%, give rules having rejection rates of about 1 per $20n$ observations. No one has explained why this should be so. No one seems to have asked.

3. The Nature of Rejection Rules

Rejection rules are not significance tests. In a study of whether, and how often, clearly spurious observations occur in a certain field, significance tests may be appropriate. But when a chemist doing routine analyses, or a surveyor making a triangulation, makes routine use of a rejection rule, he is not studying whether spurious readings occur (he may already be convinced they do sometimes), but guarding himself from their adverse effect. The same is often true of a statistician when he performs least-squares analyses, especially if he has a single very large bulk of data or many similar smaller sets of data to analyze. The statistician may feel little temptation to exercise personal judgment in "processing" the data, and prefer to follow a rigid rule for treating outliers, whose long-term effect he knows.

A rejection rule is like a householder's fire insurance policy. Three questions to be considered in choosing a policy are

(1) What is the premium?

(2) How much protection does the policy give in the event of fire?

(3) How much danger really is there of a fire?

Item (3) corresponds to the study of whether spurious readings occur in fact—a study that is hardly possible unless plenty of readings are available. The householder, satisfied that fires *do* occur, does not bother much about (3), provided the premium seems moderate and the protection good. In what currency can we express the premium charged and the protection afforded by a rejection rule? That depends on the purpose of the observations; an answer can be given as soon as a suitable loss function is specified.

Often estimation errors are only one among several types of error affecting the whole investigation or process. For example, the error made by the manufacturer in assaying a drug is not the only reason why the patient receives an incorrect dose. The doctor judges only roughly how much of the drug the patient needs, the prescription is filled by the pharmacist with less than infinite precision, and the quantity consumed by the patient may bear only a faint resemblance to the instruction of one teaspoonful every four hours. So long as the assay errors are on the whole small, there would be no interest in very detailed information about their distribution; the bias and variance will suffice. In research experiments, the immediate object may be to measure responses to certain treatments, but the ultimate object is usually to throw light on some broad class of phenomena. Estimation errors are not the only impediment to clear perception. If the investigation is in the end abortive, estimation errors will most likely not be solely responsible. Again, we do not need to know more about the estimation errors than their variance (and bias, if any).

So variance will be considered here, although in principle any other measure

of expected loss could be used. The premium payable may then be taken to be the percentage increase in the variance of estimation errors due to using the rejection rule, when in fact all the observations come from a homogeneous normal source; the protection given is the reduction in variance (or mean squared error) when spurious readings are present. Rejection rates are of no more than incidental interest.

With this approach, any suggested rejection rules can be investigated and compared. As in all previous published work on outliers, it will be assumed here that the following conditions are satisfied (they need not be in practice).

(i) Whatever circumstance causes an observation to be spurious is not expected to affect neighboring observations; all observations are supposed independent in this respect. It is also supposed that spuriousness is uncorrelated with the reading that would have been obtained had the observation been made without abnormal error.

(ii) Computation costs can be ignored. If that were not so, the "premium" would have to include the extra computation cost resulting from using the rule.

(iii) No prior knowledge concerning the means or regression coefficients that are to be estimated from the data is incorporated in the rejection rule, which is therefore "impartial." For example, suppose the observations consist of just two replicate readings, the population mean is to be estimated and the population variance is known. Any impartial rejection rule must lead to rejection if the difference between the readings exceeds some critical value, and then since there is no way of saying which is the better observation both readings must be rejected (or one after flipping a coin, perhaps). In practice, the observer may consider one of the readings reasonable and the other unlikely, and so he will retain the first and reject the other. He will then be using prior information about the population mean.

This restriction to an impartial rejection rule has an important effect on the character of our results. We return to consider it further in Section 7.

It is natural to consider rejection criteria based on the magnitude of the residual. Such a type of criterion seems likely to work better than other relatively easy criteria based on order statistics. In the least-squares analysis of complex patterns of observations, the computation of residuals is, or should be, a standard procedure, and a rejection criterion based on residuals is therefore particularly convenient. To begin with, we consider the simplest possible case, appropriate to typical chemical analyses, where only one mean is to be estimated and the population variance can be supposed known. We then pass on to complex patterns, and also consider "Studentized" criteria for use when the population variance is unknown.

Notation. Throughout, the observations will be denoted by $y_1 , y_2 , \cdots , y_n$, there being n observations in all. The residuals (before rejection of any outliers) will be denoted by $z_1 , z_2 , \cdots , z_n$. ν will always denote the number of residual degrees of freedom, i.e. ν is the rank of the matrix transforming (y_i) to (z_i).

4. Formulation of the Problem for a Simple Sample

We are given observations $y_1 , y_2 , \cdots , y_n$ $(n \geq 3)$. It is hoped they are a

random sample from a normal population $N(\mu, \sigma^2)$, where σ is known and μ is to be estimated. But possibly one or more of the y_i are spurious, coming from a different source, and ought to be rejected. We consider the effect of applying a rejection rule routinely to samples of fixed size n. Let

$$z_i = y_i - \bar{y}, \quad n\bar{y} = \sum_i y_i \qquad (i = 1, 2, \cdots, n).$$

If y_i is omitted, the average of the remaining observations is

$$\sum_{j \neq i} y_j/\nu = \bar{y} - z_i/\nu, \tag{4.1}$$

where $\nu = n - 1$. More generally, if several observations are omitted, say $y_1, y_2, \cdots, y_r$, the average of the rest is

$$\bar{y} - (z_1 + z_2 + \cdots + z_r)/(n - r). \tag{4.2}$$

Let M be the serial number of the observation having the greatest residual, so that

$$|z_M| > |z_i| \quad \text{for all} \quad i \neq M. \tag{4.3}$$

(We suppose that the observations are recorded to sufficient decimal places for no two residuals to be equal in magnitude.)

We propose to reject any observation whose residual is excessively large. The following type of rule is unsatisfactory:

RULE 0. *For given C, reject every observation y_i such that*

$$|z_i| > C\sigma.$$

Estimate μ by the mean of the retained observations.

The reason is that a single outlier, if it outlies sufficiently, can cause all the $|z|$'s to exceed $C\sigma$, and the whole sample would then be rejected. The following rule is more cautious:

RULE 1. *For given C, reject y_M if $|z_M| > C\sigma$; otherwise no rejections. Estimate μ by the mean of the retained observations, thus*

$$\hat{\mu} = \bar{y} \quad \text{if} \quad |z_M| < C\sigma,$$
$$= \bar{y} - z_M/\nu \quad \text{if} \quad |z_M| > C\sigma.$$

Under this rule, not more than one observation can be rejected. The most frequent values for n in chemical analysis are 3 and 4, and then Rule 1 is probably the best that can be suggested. If more than one observation out of a very small sample appeared to be spurious, the observer would most likely wish to scrap them all. For large samples, however, the possibility of multiple selective rejections needs to be considered, and the following rule is suggested:

RULE 2. *Apply Rule 1. If an observation is rejected, consider the remaining observations as a sample of size $n - 1$ and apply Rule 1 again; and so on. Estimate μ by the mean of the retained observations.*

It would be possible for the values of C to differ in the successive applications of Rule 1, but there is no obvious advantage in this, and, in so far as Rule 2 is considered below, C will be supposed constant. In fact, it is difficult to study Rule 2 exactly, apart from Monte Carlo computation. Rule 1 is easier, and sometimes has almost the same effect as Rule 2—namely, when there is not more than one spurious observation present in the sample, C is not very small and n is not very large.

5. Theory for Rule 1 (Simple Sample)

To study the long-run effect of application of Rule 1 to samples of size n, we consider the distribution of $\hat{\mu}$, when the y_i are interpreted as chance variables rather than as a particular realization of chance variables. The z_i and M become chance variables too.

If there are no spurious observations, the joint distribution of $(z_1, z_2, \cdots, z_n)$ is independent of the distribution of $\bar{y}$, so $\bar{y}$ and z_M are independent. Each z_i has variance $\nu\sigma^2/n$; $(n/\nu)^{\frac{1}{2}} z_i/\sigma$ has the standard normal distribution $N(0, 1)$. Let a random variable T be defined as the following function of z_M :

$$\left.\begin{aligned} T &= 0 \quad \text{if} \quad |z_M| < C\sigma, \\ &= -(n/\nu)^{\frac{1}{2}}(z_M/\sigma) \quad \text{if} \quad |z_M| > C\sigma. \end{aligned}\right\} \tag{5.1}$$

Then $\hat{\mu} = \bar{y} + \sigma T/(n\nu)^{\frac{1}{2}}$, $\bar{y}$ and T are independent, $\mathcal{E}(\hat{\mu}) = \mu$ and

$$\text{var}\,(\hat{\mu}) = \frac{\sigma^2}{n}\left(1 + \frac{\mathcal{E}(T^2)}{\nu}\right). \tag{5.2}$$

The rejection rate (proportion of observations rejected in the long run) is

$$\frac{1}{n}\,\text{ch}\,\{|z_M| > C\sigma\} = \frac{1}{n}\,\text{ch}\,\{T \neq 0\}. \tag{5.3}$$

If an observation is spurious, it is convenient to think of it as being equal to an observation from $N(\mu, \sigma^2)$ plus an extra independent error or bias. If there is one spurious observation among a sample of n, and if the bias is large enough, the observation is almost certain to be rejected under Rule 1 or Rule 2. If there are two spurious observations, and their biases are large and the difference of their biases is large, it is almost certain that both will be rejected under Rule 2.

Suppose a particular observation, y_n say, is from $N(\mu + b\sigma, \sigma^2)$, while $y_1, y_2, \cdots, y_{n-1}$ are independent from $N(\mu, \sigma^2)$. Then $\bar{y}$ is distributed $N(\mu + b\sigma/n, \sigma^2/n)$, independently of $(z_1, z_2, \cdots, z_n)$. Under Rule 1,

$$\hat{\mu} = \{\bar{y} - b\sigma/n\} + \{\sigma T/(n\nu)^{\frac{1}{2}} + b\sigma/n\}, \tag{5.4}$$

where T is defined as before by (5.1). We have

$$\mathcal{E}(\hat{\mu} - \mu)^2 = \frac{\sigma^2}{n}\left(1 + \frac{1}{\nu}\,\mathcal{E}(T + b\sqrt{\nu/n})^2\right), \tag{5.5}$$

and the rejection rate is still given by (5.3). It is to be observed that the distri-

bution of z_M, and therefore that of T, depends on b and is different from what it was before. In the absence of any rejection rule, (5.5) would be replaced by

$$\mathcal{E}(\bar{y} - \mu)^2 = \frac{\sigma^2}{n}\left(1 + \frac{b^2}{n}\right). \tag{5.6}$$

In principle, (5.2) and (5.5) can be evaluated by quadrature, though if n is higher than 4 some skill in classical n-dimensional geometry is called for. The joint distribution of the (z_i) is spherical normal in a ν-dimensional flat space. If there are no spurious observations the distribution is centered at the origin; if one (or more than one) observation has an added bias the center is away from the origin. No observation is rejected provided that

$$-C\sigma < z_i < C\sigma$$

for all i, so that the point (z_i) lies inside a region R bounded by n pairs of parallel $(\nu - 1)$-flats equidistant from the origin. For $n = 3$, $\nu = 2$, R is the interior of a regular hexagon; for $n = 4$, $\nu = 3$, R is the interior of a regular octahedron. If (z_i) lies outside R, which observation is rejected, and what sign its residual has, are determined by which one of the $2n$ faces of R is intersected by the line segment joining (z_i) to the origin. To evaluate (5.2) or (5.5) an integration must be performed over the exterior of R. For (5.2) there is great symmetry, and $\mathcal{E}(T^2)$ can be expressed, after one integration along the radius vector, as $2n$ times an integral over any one face of R. The faces are regular simplexes, and the integrand is a function only of distance from the center of the simplex. Thus for $n = 3$, $\nu = 2$, each face is a line segment, and we obtain

$$\mathcal{E}(T^2) = \frac{6}{\pi}\int_{-1/\sqrt{3}}^{1/\sqrt{3}} e^{-\frac{1}{2}C^2(1+t^2)}\left\{\frac{3}{4}C^2 + \frac{1}{1+t^2}\right\}\frac{dt}{1+t^2}.$$

For $n = 4$, $\nu = 3$, each face is an equilateral triangle and the integrand is a function of distance from its center, rather more cumbersome than the above. It is far more difficult to deal with the case of one spurious observation, since the center of the distribution no longer coincides with the center of R, and much of the symmetry is lost. Only $n = 3$ has been considered. Numerical results obtained are given below in Section 10. For further progress in the accurate evaluation of (5.2) and (5.5) it will no doubt be necessary to resort to Monte Carlo techniques.

6. Approximate Formulas

Fortunately, it seems that we do not have to wait for accurate calculations to obtain results good enough for practical use. Asymptotic approximations can be found that are easy to use and ought to be tolerably reliable, as a guide to action. Just how reliable they are can only be shown convincingly by comparing them with correct values. The comparisons in Section 10 are highly encouraging, as far as they go.

For the case of no spurious observations, the proportion of observations such that the residual exceeds $C\sigma$ in magnitude, which we shall denote by α,

is given by

$$\alpha = 2\Phi(-(n/\nu)^{\frac{1}{2}}C),$$

where

$$\Phi(t) = \int_{-\infty}^{t} \phi(u)\, du, \qquad \phi(u) = e^{-u^2/2}/\sqrt{2\pi}.$$

If t_α denotes the two-tailed α-point of the standard normal distribution, we have

$$t_\alpha = (n/\nu)^{\frac{1}{2}}C. \tag{6.1}$$

α would be exactly the rejection rate if we used Rule 0. If C is very large, α is very small. Moreover, if C is large, residuals exceeding $C\sigma$ in magnitude will tend to occur singly; it will be relatively unusual for two or more observations in the same sample of n to have residuals exceeding $C\sigma$ in magnitude. This is easily proved by noting that the joint distribution of any pair of residuals is normal with correlation coefficient different from ± 1. By taking C large enough, we can make the conditional chance that $|z_j| > C\sigma$, given that $|z_i| > C\sigma$, as small as we please, and hence the conditional chance that any one of $|z_j| > C\sigma$ for $j = 1, 2, \cdots, n$, with $j \neq i$, given that $|z_i| > C\sigma$, as small as we please.

It follows that, asymptotically as $C \to \infty$, we can ignore the possibility that more than one residual in the sample will exceed $C\sigma$ in magnitude, and we obtain

$$\begin{aligned} \mathcal{E}(T^2) &\sim n \int_{|t|>t_\alpha} t^2\phi(t)\, dt \\ &= n\{2t_\alpha\phi(t_\alpha) + \alpha\}, \end{aligned} \tag{6.2}$$

and α is asymptotically the rejection rate under Rule 1.

If C is not so large as effectively to rule out the possibility that more than one of the $|z_i|$ exceeds $C\sigma$, but nevertheless we substitute (6.2) into (5.2), it may be expected that the result will approximate the effect of Rule 2 rather than Rule 1, especially if n is fairly large. We obtain this result for the simple sample if we suppose the rejected residuals to be independent and replace the divisor $n - r$ in (4.2) by $n - 1$.

Table 1 shows $2t_\alpha\phi(t_\alpha) + \alpha$ and α as functions of t_α. The entries were readily obtained from the National Bureau of Standards table of probability functions. For comparison with the results of Section 11 it may be noted that (6.2) can be expressed alternatively in terms of the tail area of a χ^2 distribution with 3 degrees of freedom.

Suppose now that there is just one spurious observation, y_n, drawn from $N(\mu + b\sigma, \sigma^2)$, and let us suppose that b is fairly large. Then it is very probable that $M = n$. Let us assume this to be so, for the moment. Then it is easy to see that $T + (\nu/n)^{\frac{1}{2}}b$ has the following distribution: outside the interval

$$I = ((\nu/n)^{\frac{1}{2}}b - (n/\nu)^{\frac{1}{2}}C, (\nu/n)^{\frac{1}{2}}b + (n/\nu)^{\frac{1}{2}}C)$$

TABLE 1

Table of $2t_\alpha\phi(t_\alpha) + \alpha$, with α shown in brackets

t_α	0	2	4	6	8
2.9	.03826	.03630	.03442	.03263	.03092
	(.00373)	(.00350)	(.00328)	(.00308)	(.00288)
3.0	.02929	.02773	.02625	.02483	.02348
	(.00270)	(.00253)	(.00237)	(.00221)	(.00207)
3.1	.02219	.02096	.01980	.01869	.01763
	(.00194)	(.00181)	(.00169)	(.00158)	(.00147)
3.2	.01663	.01568	.01478	.01392	.01311
	(.00137)	(.00128)	(.00120)	(.00111)	(.00104)
3.3	.01234	.01161	.01091	.01026	.00964
	(.00097)	(.00090)	(.00084)	(.00078)	(.00072)
3.4	.00905	.00850	.00798	.00748	.00701
	(.00067)	(.00063)	(.00058)	(.00054)	(.00050)
3.5	.00657	.00616	.00577	.00540	.00505
	(.00047)	(.00043)	(.00040)	(.00037)	(.00034)
3.6	.00472	.00442	.00413	.00385	.00360
	(.00032)	(.00029)	(.00027)	(.00025)	(.00023)

the distribution is that of a standard normal variable $N(0, 1)$; inside I the chance is concentrated at the mid-point $(\nu/n)^{\frac{1}{2}}b$. We therefore have

$$\mathcal{E}(T + (\nu/n)^{\frac{1}{2}}b)^2 = 1 + \int_I (\nu b^2/n - t^2)\phi(t)\,dt.$$

For b positive and large, we can replace the upper end-point of I by ∞, obtaining

$$\mathcal{E}(T + (\nu/n)^{\frac{1}{2}}b)^2 = 1 + (\nu b^2/n - 1)\Phi(-x) - x\phi(x), \tag{6.3}$$

where

$$x = (\nu/n)^{\frac{1}{2}}b - (n/\nu)^{\frac{1}{2}}C.$$

If (6.3) is substituted into the right-hand side of (5.5), the result will be a lower bound for $\mathcal{E}(\hat{\mu} - \mu)^2$, because we have ignored the possibility that M may be different from n. We may hope, however, that the bound will be close enough to the true value to indicate how large b needs to be for the rejection rule to give good protection.

7. Complex Patterns of Data

Let us turn now to complex patterns of data. We suppose the observations $y_1, y_2, \cdots, y_n$ (if none is spurious) to be drawn from independent normal distributions having common variance σ^2 and means that are given linear functions of some unknown parameters. At first we shall suppose σ^2 known.

When all the unknown parameters have been estimated by least squares, the residuals $\{z_i\}$ can be calculated.

We observe first of all that in general the residuals do not all have the same chance distribution. This is illustrated by simple linear regression: y_i is drawn from $N(\mu + \beta x_i, \sigma^2)$, where the x_i are predetermined, and let us say $\Sigma_i x_i = 0$. Then

$$z_i = y_i - \bar{y} - (\Sigma_i y_i x_i / \Sigma_i x_i^2) x_i ,$$

and we find

$$\text{var}\,(z_i) = \left(\frac{n-1}{n} - \frac{x_i^2}{\Sigma_i x_i^2}\right)\sigma^2 .$$

Thus $0 \leq \text{var}\,(z_i) \leq (n-1)\sigma^2/n$, both bounds being attainable for some i and some set of $\{x_i\}$. If the $\{x_i\}$ are equal-spaced, we have

$$(n-4)\sigma^2/n < \text{var}\,(z_i) \leq (n-1)\sigma^2/n,$$

and if n is not very small we shall not go far wrong if we suppose all the residuals to have the same variance, say the average value over i, which is $(n-2)\sigma^2/n$, or $\nu\sigma^2/n$. All the residuals have exactly this variance if (and only if) there are just two different x-values, with $n/2$ observations at each.

In order to keep our considerations as simple as possible, we shall restrict attention to cases where (in the absence of spurious readings) all residuals have exactly the same variance, namely $\nu\sigma^2/n$. All ordinary factorial designs, where the different levels of each factor are replicated equally often, have this property, and also Latin squares and balanced incomplete block designs. A type of design to be excluded is a factorial design where the levels of one of the factors are not equally replicated, as in an agricultural experiment on insecticides, where it would usually be advisable to replicate the "control" treatment of no insecticide more heavily than any one of the alternative insecticide treatments under test. It is to be expected that results obtained under the assumption of equal variances will be approximately correct if the residuals have only approximately equal variances, but we shall not examine this question.

We observe next that in general not every pair of residuals has the same correlation. For some designs, some of the correlation coefficients are ± 1. An example is the 3×3 Latin square, where the residuals are equal in sets of 3, according to the letters of the orthogonal square, and so the correlation coefficients between the 36 possible pairs of the 9 residuals are 9 of them equal to 1 and 27 of them equal to $-\frac{1}{2}$. Suppose we are given that $\sigma = 1$ and the lay-out and yields are as follows, where *0*, *1*, *2* denote the three levels of a treatment:

(*0*)	$y_1 = 13.9$	(*1*)	$y_2 = 5.9$	(*2*)	$y_3 = 6.3$
(*2*)	$y_4 = 6.0$	(*0*)	$y_5 = 5.7$	(*1*)	$y_6 = 6.4$
(*1*)	$y_7 = 6.0$	(*2*)	$y_8 = 6.3$	(*0*)	$y_9 = 4.9$

The residuals are

$$z_1 = 2.04 \quad z_2 = -1.22 \quad z_3 = -0.82$$
$$z_4 = -1.22 \quad z_5 = -0.82 \quad z_6 = 2.04$$
$$z_7 = -0.82 \quad z_8 = 2.04 \quad z_9 = -1.22$$

Since the standard deviation of residuals is $(\nu/n)^{\frac{1}{2}}\sigma = 0.47$, the values 2.04 may be judged excessively large and indicative of a spurious reading. The problem is, which reading? If we reject any *one* of y_1, y_6 and y_8 and treat that observation as missing, the estimated value to be substituted is equal to the observed value minus n/ν $(= 4.5)$ times the residual (2.04), i.e. we subtract 9.2 from *one* of y_1, y_6, y_8, and proceed to estimate row, column and treatment effects in the usual way. The new residuals are

$$z_1 = 0 \quad z_2 = -0.20 \quad z_3 = 0.20$$
$$z_4 = -0.20 \quad z_5 = 0.20 \quad z_6 = 0$$
$$z_7 = 0.20 \quad z_8 = 0 \quad z_9 = -0.20$$

and these are satisfactorily small. But our estimates of the row, column and treatment effects will depend sharply on which one of the three questionable observations is rejected. If we insist on an impartial rejection rule, not using any prior knowledge concerning the unknown parameters to be estimated, we cannot have any preference, and to settle the deadlock we might consider rejecting all three observations. We should then scrap the whole experiment, because (in the absence of prior knowledge) too few readings are left for the unknown parameters to be uniquely estimated.

In practice the experimenter always has some prior knowledge, which may be small compared with good observations having $\sigma = 1$ but is perhaps appreciable compared with a gross error, here estimated at 9.2. If he rejects y_6 or y_8, the fitted yield will be negative, which he may find disturbing. Apart from that, he very likely thinks that rather small row, column and treatment effects are more probable than large ones, and therefore the estimates obtained after rejecting y_1 are more plausible than those obtained after rejecting either y_6 or y_8. In fact, he may guess that just the initial digit of y_1 is spurious, and y_1 should read 3.9.

One may conclude from this example that it is unwise to try to apply an impartial routine rejection rule to data such that some pairs of residuals have correlation ± 1. It may be wise to go further and exclude designs such that some pairs of residuals have large correlations, say greater than $\frac{1}{2}$ or $\frac{3}{4}$ in magnitude, because if a spurious reading occurs it is not unlikely that another reading will be rejected instead. Thus it is of interest to know, in respect of any proposed design, what correlations the residuals will have, and this is a matter not discussed in books on the design of experiments.

Let the observations $\{y_i\}$ be represented by the $n \times 1$ column vector $\mathbf{y}$, let $\boldsymbol{\theta}$ be a $(n - \nu) \times 1$ vector of nonredundant unknown parameters and let $\mathbf{A}$

be the $n \times (n - \nu)$ matrix of known coefficients, such that, in the absence of spurious readings,

$$\mathbf{y} \text{ has a spherical normal distribution with } \mathcal{E}(\mathbf{y}) = \mathbf{A}\boldsymbol{\theta}. \tag{7.1}$$

$\mathbf{A}$ is of rank $n - \nu$, and $\mathbf{V}$ exists such that $\mathbf{A}'\mathbf{A} = \mathbf{V}^{-1}$. The $n \times 1$ vector of residuals $\mathbf{z}$, after least-squares estimation of $\boldsymbol{\theta}$, is given by

$$\mathbf{z} = \mathbf{Q}\mathbf{y}, \tag{7.2}$$

where $\mathbf{Q} = \mathbf{I}_n - \mathbf{AVA}'$, $\mathbf{I}_n$ being the $n \times n$ unit matrix. $\mathbf{Q}$ has the following properties:

(i) $\mathbf{Q}$ is a $n \times n$ symmetric idempotent matrix of rank ν. It correctly transforms $\mathbf{y}$ to $\mathbf{z}$, when $\boldsymbol{\theta}$ is estimated by least squares *as if* (7.1) were true, whether or not (7.1) is true in fact. It is invariant to transformations of the parameter set $\boldsymbol{\theta}$.

(ii) The ith row of $\mathbf{Q}$ shows how a gross error in y_i is distributed among the residuals (as we see by replacing $\mathbf{y}$ by a vector consisting of 1 in the ith place and 0's elsewhere).

(iii) The variance matrix of the residuals, if (7.1) is true, is given by

$$\mathcal{E}(\mathbf{z}\mathbf{z}') = \mathbf{Q}\sigma^2. \tag{7.3}$$

Two practical methods of calculating $\mathbf{Q}$ are:-

(i) Solve the least-squares equations in any convenient manner and express the $\{z_i\}$ in terms of the $\{y_i\}$. For sufficiently symmetrical designs, this need be done for only one z, say z_1 .

(ii) If it is *convenient* (it is always *possible*) to choose $\boldsymbol{\theta}$ so that $\mathbf{V}$ is a scalar multiple of $\mathbf{I}_{n-\nu}$, say $\mathbf{V} = v\mathbf{I}_{n-\nu}$, then $\mathbf{Q} = \mathbf{I}_n - v\mathbf{AA}'$, and to find $\mathbf{AA}'$ we calculate scalar products of pairs of rows of $\mathbf{A}$. This method is easy for 2^k factorial designs—for which, however, it is almost the same as (i).

If all residuals have the same variance, all coefficients in the principal diagonal of $\mathbf{Q}$ are equal to ν/n, and from the relation $\mathbf{Q}^2 = \mathbf{Q}$ we easily deduce that the mean squared correlation coefficient between pairs of residuals is

$$\overline{\rho^2} = \frac{n - \nu}{(n - 1)\nu}. \tag{7.4}$$

Thus the square root of this quantity is a lower bound to the magnitude of the largest correlation coefficient between any pair of residuals. Only rarely (i.e. for a small proportion of available designs) is this lower bound attained.

If the observations have a simple two-way classification, in k rows and l columns, with $n = kl$, and the expectation of each observation y is a row constant plus a column constant, the possible correlations between residuals are easily seen to be

$$-\frac{1}{\kappa}, -\frac{1}{\lambda}, \frac{1}{\kappa\lambda}, \tag{7.5}$$

where $\kappa = k - 1, \lambda = l - 1$. If either k or $l = 2$, there are correlations of -1.

If the observations have a one-way classification, say k observations in each of l columns, so that the expectation of each observation is just a column constant, the possible correlations between residuals are

$$-\frac{1}{\kappa}, 0; \tag{7.6}$$

again we have correlations of -1 if $k = 2$. In particular, there will be correlations of -1 if the design consists of two replications, in either one block or two blocks, of some factorial design and if all interactions between the factors are estimated, as well as the general mean and main effects and (for two blocks) the block difference.

Table 2 lists the correlations that occur among the residuals for some designs with fairly small n. Designs having correlations of ± 1 have been omitted, except for designs (2) and (14), which have been included for comparative purposes. The correlations for design (2) are an example of (7.6); those for design (3) an example of (7.5). In designs (1) and (2), three two-level factors are compared in eight observations. Both designs can be obtained by deleting four factors from a "saturated" $2^7/16$ design, as given by Plackett and Burman (1946); and it is interesting to note that two essentially different designs can be obtained in this way. Design (4) is obtained by deleting any six factors from the Plackett-Burman saturated fraction of 2^{11}. Designs (7) and (8) are half replicates of a 2^5 factorial. (7) is the obvious choice, and achieves equal-magnitude correlations. (8) illustrates the effect of another alias subgroup. The design "$2^5/2$: ABC" has correlations differing from those of (8) in sign only. Designs (9), (11) and (14) were given by Brownlee, Kelly and Loraine (1948) and have often been quoted since, as the "optimum" fractions, such that main effects are not confounded with any two-factor interactions. With respect to correlations among residuals they are not optimum, and are bettered by designs (10), (13) and (15), which, with (16) and (17), have alias subgroups permitting the lowest possible largest correlation. Design (12) is of interest, because it can be reinterpreted as two replications (in two blocks) of a 2^3 factorial, with all two-factor interactions estimated. If the three-factor interaction is also estimated, we have at once the correlation pattern of (7.5), with some correlations equal to -1. Designs (16) and (18) are essentially the same design; like (7), they achieve equal-magnitude correlations. Design (24) can be reinterpreted as a 3^3 factorial with all two-factor interactions estimated. Of all these twenty-four listed designs, (4) alone is unsymmetrical, in the sense that all the rows of $\mathbf{Q}$ do not contain exactly the same set of coefficients.

8. Formulation for Complex Patterns

Restricting attention, then, to designs such that, in the absence of spurious observations, all residuals have the same variance and no two residuals have correlation coefficient equal, or very close, to ± 1, we wish to examine the effect of routine application of a rejection rule. Two questions to consider first are: what sort of rule, and effect on what?

Let M be defined as before, by (4.3). Rules 0 and 1 of Section 4 can be adapted

TABLE 2

Correlations between residuals for various designs

Design	n	v	$\sqrt{\frac{n-v}{(n-1)v}}$	Correlations between residuals
(1) 2^3	8	4	0.378	$\pm\frac{1}{3}$, 0
(2) $2^3/2$ repeated: *ABC*	8	4	0.378	-1, 0
(3) 3^2	9	4	0.395	$-\frac{1}{2}$, $\frac{1}{4}$
(4) Plackett-Burman fraction of 2^5	12	6	0.302	$\pm\frac{1}{5}$, $\pm\frac{1}{5}$, 0
(5) Balanced incomplete blocks ($v = 4$, $k = 3$)	12	5	0.357	$-\frac{1}{2}$, $\frac{1}{4}$, $-\frac{1}{8}$, $\frac{1}{10}$
(6) 2^4	16	11	0.174	$\pm\frac{3}{11}$, $\pm\frac{1}{11}$
(7) $2^5/2$: *ABCDE*	16	10	0.200	$\pm\frac{1}{3}$
(8) $2^5/2$: *ABCD*	16	10	0.200	$\pm\frac{1}{5}$, $\pm\frac{1}{5}$, 0
(9) $2^6/4$: *ABCD, CDEF*	16	9	0.228	$\frac{5}{9}$, $-\frac{1}{9}$, $\pm\frac{1}{9}$
(10) $2^6/4$: *ABC, DEF*	16	9	0.228	$-\frac{1}{3}$, $\frac{1}{9}$
(11) $2^7/8$: *ABCD, CDEF, ACEG*	16	8	0.258	$\frac{1}{4}$, $-\frac{1}{4}$, 0
(12) $2^7/8$: *BCE, CDF, BDG*	16	8	0.258	$-\frac{1}{4}$, $\pm\frac{1}{4}$, 0
(13) $2^7/8$: *ABC, CDE, EFG*	16	8	0.258	$-\frac{1}{3}$, $\pm\frac{1}{4}$, 0
(14) $2^8/16$: *ABCD, CDEF, ACEG, EFGH*	16	7	0.293	1, $-\frac{1}{7}$
(15) $2^8/16$: *ABC, CDE, EFG, AGH*	16	7	0.293	$\pm\frac{3}{7}$, $\pm\frac{1}{7}$
(16) $2^9/32$: *ABC, CDE, EFG, AGH, BFI*	16	6	0.333	$\pm\frac{1}{3}$
(17) $2^{10}/64$: *ABE, ACF, ADG, BCH, BDI, CDJ*	16	5	0.383	$-\frac{3}{5}$, $\frac{1}{5}$
(18) Latin square	16	6	0.333	$\pm\frac{1}{3}$
(19) Balanced incomplete blocks ($v = 7$, $k = 3$)	21	8	0.285	$-\frac{1}{3}$, $\frac{1}{4}$, $-\frac{1}{3}$
(20) Latin square	25	12	0.212	$-\frac{1}{4}$, $\frac{1}{8}$
(21) Graeco-Latin square	25	8	0.298	$\frac{1}{3}$, $-\frac{1}{4}$
(22) 3^3	27	20	0.116	$-\frac{1}{5}$, $\frac{1}{10}$, $-\frac{1}{20}$
(23) $3^4/3$: *ABCD*	27	18	0.139	$\pm\frac{1}{4}$, 0
(24) $3^9/729$: *ABD, AB²E, ACF, AC²G BCH, BC²I*	27	8	0.302	$-\frac{1}{3}$, $\frac{1}{4}$, $-\frac{1}{8}$

Most of the designs are factorials. Thus (10) is a quarter replicate of a factorial arrangement with 6 factors *A*, *B*, *C*, *D*, *E*, *F*, each at 2 levels; *ABC* and *DEF* are generators of the alias subgroup. (2) is a half replicate repeated in one block. (4) is given by Plackett and Burman (1946). In the balanced incomplete block designs (5) and (19), v is the number of varieties or treatment levels, k is the number of units in each block. *It is to be understood that no interactions are estimated.* Several designs can be reinterpreted as fewer-factor designs with two-factor interactions estimated, as mentioned in the text. For the incomplete block designs, interblock information is supposed not recovered.

to the present case, by changing the second sentence to read: *Estimate the unknown parameters from the retained observations by the method of least squares.* Normally it will be appropriate to perform this estimation by the two-stage method known misleadingly as "analysis of covariance," the rejected observations being regarded as missing values to be estimated.

Suppose a particular observation, y_1 say, is missing or (if present) is held

to be spurious. Let $\{z_i\}$ be the residuals when some arbitrary value for y_1 is used in the least squares estimation of the parameters. Then if y_1 is replaced by $y_1 - \eta$, the residuals are replaced by $\{z_i^{(1)}\}$, where

$$z_i^{(1)} = z_i - \eta q_{i1} , \tag{8.1}$$

$\{q_{ij}\}$ being the coefficients of $\mathbf{Q}$. The value of η which minimizes $\Sigma_i (z_i^{(1)})^2$ is

$$\hat{\eta} = \Sigma_i z_i q_{i1} / \Sigma_i q_{i1}^2 .$$

Given our assumptions, we have $\Sigma_i q_{i1}^2 = q_{11} = \nu/n$ and $\Sigma_i z_i q_{i1} = \Sigma_{ij} q_{i1} q_{ij} y_j = \Sigma_j q_{1j} y_j = z_1$. Thus

$$\hat{\eta} = (n/\nu) z_1 . \tag{8.2}$$

If all observations except y_1 are good, the correct least-squares estimates of the unknown parameters are found by substituting $y_1 - \hat{\eta}$ for y_1 in the expressions that would have been valid if all observations, including y_1 , had been good. The residuals are now given by (8.1) with (8.2) substituted, that is

$$z_1^{(1)} = 0, \qquad z_i^{(1)} = z_i - \rho_{i1} z_1 \qquad (i \neq 1),$$

where ρ_{i1} is the correlation between z_i and z_1 ; and we obtain

$$\text{var}\,(z_i^{(1)}) = (1 - \rho_{i1}^2) \nu \sigma^2 / n. \tag{8.3}$$

Thus Rule 1 (which is the only rule that we consider in detail) prescribes that if $|z_M| > C\sigma$, y_M must be replaced by $y_M - n z_M/\nu$ in the parameter estimates.

How should our former Rule 2 be adapted? We have just seen that only for some rare designs are all the correlations ρ_{ij} equal in magnitude. (8.3) shows that if one observation is rejected, the new residuals will have unequal variance, in general. Provided the correlations are fairly small, it may seem reasonable—it is certainly simplest—to take no account of the changes in variance in formulating the rule, which would run: *Apply Rule 1. If an observation is rejected, compute revised residuals and apply Rule 1 again; and so on. Finally compute the least-squares estimates of the unknown parameters from the retained observations.*

Consider now how to measure the effect of routine application of one of these rules. We have argued in Section 3 above that it is appropriate to consider the sampling variances of estimates. Which estimates? In a complex design we are not necessarily interested in all the parameters that have to be estimated. The general mean and the block differences (if any) are usually of no more than secondary interest; that is, the experiment cannot be said to have been performed *in order to* estimate them. Sometimes the same is true even of the responses to some of the treatment factors. Let us therefore divide the parameters into two sets, those that are "of interest" and the remainder, and propose to study the variance matrix of the estimates of the parameters "of interest." (We shall impose a mild restriction on the choice of parameters for the "of interest" set.) Let the number of parameters "of interest" be κ. By an orthogonal transformation, the covariances of the parameter estimates can be made zero, and the product of the variances after transformation is equal

to the determinant of the variance matrix before transformation. A possible measure of the "premium" charged by a rejection rule is the percentage increase in the determinant of the variance matrix caused by using the rule, when in fact no observations are spurious. We shall see that, for Rule 1 with given C, this percentage increase is proportional to κ. Rejecting an observation implies a loss of information concerning each of the parameters (in the absence of spurious observations), and in a sense it is true that the greater κ is the greater is the loss.

However, to aggregate the losses for each of the κ parameters is not altogether reasonable as a measure of the premium exacted by the rule—as we see by considering the grouping of observations. It is commonly held that to group normally distributed observations at a grouping interval not exceeding $\sigma/2$ has a negligible effect on the estimation of means, and is therefore to be recommended for ease of computation, whenever ease matters. Such a grouping increases the residual variance by about 2%, and so the sampling variance of every estimated parameter by the same percentage. If the reader considers that the recommendation about grouping is equally sound however many parameters are to be estimated, he will presumably consider the following definition of the premium charged by a rejection rule to be reasonable: *the premium is that proportional increase in σ^2 which would inflate the determinant of the variance matrix of the estimated parameters of interest by as much as the rejection rule does, when no observations are spurious.*

9. Theory for Rule 1 (Complex Patterns)

We are supposing that (in the absence of spurious readings and with no rejection rule applied)

(i) all residuals have the same variance, and

(ii) no two residuals have correlation coefficient equal or close to ± 1.

We shall also suppose that

(iii) if just the κ parameters of interest are estimated, the remainder being set equal to 0, the resulting residuals all have equal variance.

These conditions concern the design—specifically, they condition the linear space spanned by the columns of $\mathbf{A}$—and they also concern the choice of the κ parameters of interest out of the total of $n - \nu$ parameters. Despite references to variance and correlations, they are independent of distribution assumptions. Condition (iii) implies, for example, that we are not allowed to pick out as the "of interest" set a one-degree-of-freedom component, e.g. the linear component, from the main effect of a three-level factor.

Let the parameters of interest be the first κ components of $\boldsymbol{\theta}$. We wish to find the ratio of determinants of the variance matrices of the estimates $(\hat{\theta}_1, \hat{\theta}_2, \cdots, \hat{\theta}_\kappa)$ calculated under the assumptions (a) that Rule 1 is applied, (b) that Rule 1 is not applied—or that Rule 1 is applied with $C = \infty$. We observe first of all that Rule 1 does not depend on the parameterization. If a nonsingular lower-triangular transformation $\mathbf{T}$ is applied to $\boldsymbol{\theta}$ and the problem is rephrased in terms of $\boldsymbol{\theta}^* = \mathbf{T}\boldsymbol{\theta}$, no change is made in the residuals or in the operation of

the rejection rule, and $\mathbf{T}$ transforms the estimates of $\boldsymbol{\theta}$ to the estimates of $\boldsymbol{\theta}^*$, i.e. $\hat{\boldsymbol{\theta}}^* = \mathbf{T}\hat{\boldsymbol{\theta}}$. In particular, the leading κ-rowed minor of $\mathbf{T}$ transforms $(\hat{\theta}_1, \hat{\theta}_2, \cdots, \hat{\theta}_\kappa)$ to the corresponding estimates of interest $(\hat{\theta}_1^*, \hat{\theta}_2^*, \cdots, \hat{\theta}_\kappa^*)$. It follows that the ratio of determinants that we seek is the same for the latter estimates as for the former. There is no loss of generality in supposing that the parameters $\boldsymbol{\theta}$ satisfy the conditions

(iv) the parameters of interest are the first κ components of $\boldsymbol{\theta}$,

(v) $$\mathbf{V} = \mathbf{I}_{n-\nu}/n.$$

(If (v) is not already satisfied, apply the lower-triangular transformation $\mathbf{T}$ to $\boldsymbol{\theta}$ defined by $n\mathbf{T}'\mathbf{T} = \mathbf{V}^{-1}$.)

From (i) and (v) it easily follows that the sum of squares of the coefficients in any row of $\mathbf{A}$ is equal to $n - \nu$; (v) also implies that the columns of $\mathbf{A}$ are orthogonal and the sum of squares of coefficients in any column is equal to n. If we now introduce condition (iii), it is easy to deduce that the sum of squares of the first κ coefficients in any row of $\mathbf{A}$ is equal to κ. In the absence of any rejection rule, the least-squares estimates of $\boldsymbol{\theta}$ are

$$\hat{\boldsymbol{\theta}} = \mathbf{A}'\mathbf{y}/n.$$

Rule 1 has the effect of replacing y_M by $y_M + (n/\nu)^{\frac{1}{2}}\sigma T$, where T is defined by (5.1). Let us consider distributions conditionally on the value of M. We can apply an orthogonal transformation to the parameters of interest, so that the transformed $\mathbf{A}$ satisfies

$$a_{M2} = a_{M3} = \cdots = a_{M\kappa} = 0,$$

and then, because the transformation does not disturb the sum of squares of the first κ coefficients in any row of $\mathbf{A}$, we must have

$$a_{M1} = \sqrt{\kappa}.$$

(The sign is arbitrary; let us take the positive root.) The rejection rule does not now affect $\hat{\theta}_2, \hat{\theta}_3, \cdots, \hat{\theta}_\kappa$, but $\hat{\theta}_1$ is modified by the rejection rule by the addition of the term

$$(\kappa/n\nu)^{\frac{1}{2}}\sigma T.$$

If in fact no observations are spurious, this added term is independent of what it is added to (which has variance σ^2/n), and so we get the following result. We can transform the parameters of interest, by an orthogonal transformation depending on M, so that one new parameter has its variance changed to

$$(\sigma^2/n)\{1 + (\kappa/\nu)\mathcal{E}(T^2)\},$$

while all the other components have unaltered variances of σ^2/n. The determinant of the variance matrix (which is invariant to orthogonal transformation of parameters) is thus

$$(\sigma^2/n)^\kappa\{1 + (\kappa/\nu)\mathcal{E}(T^2)\}. \tag{9.1}$$

This result is independent of M and therefore true unconditionally. The factor

in curly brackets is the ratio that we are looking for, and is invariant to any nonsingular transformation of the parameters. The premium, according to the suggested definition at the end of Section 8, is

$$\{1 + (\kappa/\nu)\mathcal{E}(T^2)\}^{1/\kappa} - 1, \tag{9.2}$$

which is less than, but may be closely equal to,

$$(1/\nu)\mathcal{E}(T^2), \tag{9.3}$$

unless $\kappa = 1$, when these expressions are exactly equal.

Now suppose that one observation, say y_n, is spurious and has bias $b\sigma$. It may happen that $M = n$, or, if not, that the first κ coefficients in the Mth row of $\mathbf{A}$ are the same as those in the nth row. Conditionally on this being so, our previous line of reasoning shows that the determinant of the matrix of mean squared estimation errors is changed from (9.1) to

$$(\sigma^2/n)^\kappa\{1 + (\kappa/\nu)\mathcal{E}(T + (\nu/n)^{\frac{1}{2}}b)^2\}. \tag{9.4}$$

This result does not hold unconditionally, but suffices for application of the results of Section 6. If no rejection rule were applied, (9.4) would be replaced by

$$(\sigma^2/n)^\kappa\{1 + (\kappa/n)b^2\}. \tag{9.5}$$

The arguments of Section 6, used to derive (6.2) and (6.3), are still valid in this context. We are now in a position to draw some conclusions.

10. Numerical Results

Both for the simple sample and for the fairly broad type of complex patterns just considered, we see, from (6.2) and (5.2) or (9.3), that the premium charged by Rule 1 can be roughly reckoned at

$$(n/\nu)\{2t_\alpha\phi(t_\alpha) + \alpha\}, \tag{10.1}$$

where $t_\alpha = (n/\nu)^{\frac{1}{2}}C$, as explained in Section 6. This should be tolerably accurate if C is so large that $n\alpha$ is quite small. If C is not so large, there is some reason to expect that (10.1) approximates the premium charged by Rule 2 rather than Rule 1.

How much premium we are willing to pay should depend on how greatly we fear spurious observations. But, as with domestic insurance, we shall probably not care very much provided the premium is small. Let us see what can be had for a premium of 2%, i.e. an effective increase of 2% in the residual variance. Setting the expression (10.1) equal to 0.02, we obtain the values of C and α given in the first part of Table 3. The lower part of the table refers to a premium of 1%. The value 1 for ν/n is not attainable, but a simple sample of large size has ν/n close to 1. The least value of ν/n for the designs given in Table 2 was 8/27, or about 0.3. Lower values for ν/n are no doubt rare, except perhaps in some screening experiments.

The protection given by Rule 1, when one observation is spurious with a large bias $b\sigma$, is given approximately by substituting (6.3) into (5.5) or (9.4) and comparing the result with (5.6) or (9.5), respectively. The best the rule

could do would be always to reject the spurious observation, and in that case we should have

$$\mathcal{E}(T + (\nu/n)^{\frac{1}{2}}b)^2 = 1.$$

This is the limit of (6.3) as $b \to \infty$. To give some idea of how large b must be for the rule to work well, values of b are tabulated for which, according to (6.3),

$$\mathcal{E}(T + (\nu/n)^{\frac{1}{2}}b)^2 = 1.5.$$

The chance that the spurious observation will escape rejection is, in the cases tabulated, of the order of 0.02. For somewhat smaller b the rejection rule will be beneficial, better than no rejection rule, but the mean squared estimation error will be substantially greater than if there had been no spurious observation. If $b = (n/\nu)C$, the chance that the spurious observation will be rejected is about 0.5.

For given n and no spurious observations, as ν is reduced the effect of a rejection becomes more serious. Consequently the rejection rate α must be reduced to preserve a fixed premium, and $(n/\nu)^{\frac{1}{2}}C$ is increased. On the other hand, for a fixed bias b, $(\nu/n)^{\frac{1}{2}}b$ is reduced, and so the x in (6.3) is reduced twice over. Hence the striking increase in the b's as one reads across Table 3. If we increase the premium a little, the b's are not greatly reduced.

Thus for an experiment of fixed size n, as the design is made more ingenious and ν becomes smaller, a gross error in one of the readings spoils more estimated effects and becomes less detectable. Even if we can prevent the correlations from blowing up to ± 1, it takes a bigger error to be seen.

The above findings indicate that, while the rejection rate α should depend on ν/n, it should not otherwise depend upon n. To that extent they support Wright against most other authors, who have advocated fixed rejection rates per experiment.

It remains to compare the approximate formulas with exact calculations. Com-

TABLE 3

Tabulation of rejection rules, based on approximate formulas

ν/n	1.0	0.8	0.6	0.4	0.2
2% premium					
C	3.14	2.87	2.56	2.18	1.63
α	0.00171	0.00131	0.00094	0.00058	0.00026
b	5.1	5.8	6.9	8.7	12.8
1% premium					
C	3.37	3.08	2.73	2.31	1.72
α	0.00076	0.00058	0.00042	0.00026	0.00012
b	5.4	6.1	7.2	9.1	13.3

For given premium, values of C and α are given, and also that value of b such that a bias of size $b\sigma$ in one observation inflates the mean squared estimation error only one and a half times as much as rejection of one observation at random when all observations are good.

TABLE 4

Tabulation of rejection rules, based on accurate calculations

C	(n/σ^2) var (Δ)	Rejection rate	Approximate (n/σ^2) var (Δ)	Approximate rejection rate
Simple sample, $n = 3$, $\nu = 2$				
2.46003	1.04	0.002433	1.04241	0.002588
2.66184	1.02	0.001065	1.02088	0.001114
2.84623	1.01	0.000475	1.01032	0.000490
3.01724	1.005	0.000214	1.00512	0.000220
Simple sample, $n = 4$, $\nu = 3$				
2.57994	1.04		1.04134	0.002891
2.79541	1.02		1.02043	0.001247
2.99206	1.01		1.01014	0.000550
3.17434	1.005		1.00504	0.000247

The table refers to Rule 1 when there are no spurious observations, and shows values of C corresponding to given premiums of 4%, 2%, 1%, and ½%, together with the rejection rate. The last two columns show for comparison values obtained by the approximate formulas, using the C values quoted.

[Results for $n = 3$, $\nu = 2$, one spurious observation, are to follow.]

putations for the simple sample with $n = 3$, Rule 1, were begun by Mr. William W. Hardgrave, using a desk machine. Later, one of us (I. G.) used an IBM 650 to check and extend the calculations. Some results are presented in Table 4.

11. σ UNKNOWN

Clearly the above investigation leaves many questions unanswered. Further calculations are needed to determine the reliability of the approximate formulas. One would like to know the effect of routine application of a rejection rule to data such that the component of random variation was homogeneous but not normal—for example, having a distribution of the same shape as Student's distribution with 7 degrees of freedom (Pearson's Type VII with exponent 4), as suggested by Jeffreys, or having a "contaminated" distribution of the type considered by Tukey (1960). How does the rule function in the presence of Tukey's (1949) removable nonadditivity, or of other systematic departures from standard assumptions?

In view of statistical history during the past fifty years, the question likely to be asked first is: what happens if σ is unknown? It is not harder in principle to investigate Studentized rejection rules than those already considered. Let s^2 denote the best (quadratic unbiased) estimate of σ^2 available. If derived from the given observations (y_i) only, s^2 has ν degrees of freedom. But in general we may suppose that there is also some prior information about σ^2, equivalent to a quadratic estimate having ν_0 degrees of freedom, so that s^2 has $\nu + \nu_0$ degrees of freedom. The condition $|z_M| > C\sigma$ in Rule 1 and in the definition of T, (5.1), will be replaced by $|z_M| > Cs$.

For the case of no spurious observations, we can obtain an approximate expression for the premium, corresponding to (10.1), as follows. Choosing any one of the residuals, z_1 say, for consideration, we can make a fixed orthogonal change of axes so that (z_i) is transformed to $(\zeta_1, \zeta_2, \cdots, \zeta_\nu, 0, \cdots, 0)$, where where the ζ's have independent $N(0, \sigma^2)$ distributions, and $\zeta_1 = (n/\nu)^{\frac{1}{2}} z_1$. It follows that the condition $|z_1| > Cs$ is equivalent to

$$\left(\frac{\nu(\nu + \nu_0)}{nC^2} - 1\right)\zeta_1^2 > \sigma^2\chi^2, \tag{11.1}$$

where χ^2 stands for a χ^2 variable with $\nu + \nu_0 - 1$ degrees of freedom, independent of ζ_1. The chance α that this condition is satisfied, which is approximately the rejection rate under Rule 1, is given by

$$\sqrt{\frac{nC^2}{\nu}\,\frac{\nu + \nu_0 - 1}{\nu + \nu_0 - nC^2/\nu}} = t_\alpha^{(\nu+\nu_0-1)}, \tag{11.2}$$

where the right-hand side denotes the two-tailed α-point of Student's distribution with $\nu + \nu_0 - 1$ degrees of freedom. This formula corresponds to (6.1) and extends a result given by Thompson. It can be expressed alternatively in terms of the incomplete beta function ratio tabulated by Karl Pearson, thus:

$$\alpha = I_x\left(\frac{\nu + \nu_0 - 1}{2}, \frac{1}{2}\right),$$

where $x = 1 - nC^2/\nu(\nu + \nu_0)$. The premium $\nu^{-1}\mathcal{E}(T^2)$ is given approximately by $(n/\nu\sigma^2)$ times the partial expectation of ζ_1^2, integration being confined to the region where (11.1) is satisfied. We obtain the result:

$$\frac{1}{\nu}\,\mathcal{E}(T^2) \sim \frac{n}{\nu} I_x\left(\frac{\nu + \nu_0 - 1}{2}, \frac{3}{2}\right), \tag{11.3}$$

where x is as above. This may be expressed alternatively in the following rule for determining C, given the premium. Multiply the premium by ν/n, and find the corresponding upper percentage point of the variance ratio (F) distribution with 3 and $\nu + \nu_0 - 1$ degrees of freedom. Calling this F, we have

$$\frac{nC^2}{\nu} \sim \frac{3F}{1 + (3F - 1)/(\nu + \nu_0)}.$$

For example, suppose that $\nu/n = 0.5$, $\nu + \nu_0 = 30$ and the premium is to be 0.02. We find that the upper 1% (= 0.02 × 0.5) point of F with 3 and 29 degrees of freedom is 4.54 and so

$$2C^2 = \frac{13.6}{1.42} = 9.59, \qquad C = 2.19,$$

and from (11.2) we have $\alpha = 0.00092$. If $\nu + \nu_0 = 121$, the premium and ν/n remaining the same, we find similarly $C = 2.33$, $\alpha = 0.00079$; and for $\nu + \nu_0 = \infty$ (σ known) we have $C = 2.38$, $\alpha = 0.00076$. For a given premium, C depends quite sharply on $\nu + \nu_0$, and to a lesser extent α also varies with $\nu + \nu_0$.

For such numerical calculations, Federighi's table (1959) is useful for (11.2)

and Table 16 in the Biometrika Tables (1954) for (11.3). Ingenious interpolation is called for.

It would be possible to investigate the response of the rejection rule to a spurious observation, as in Section 6, but I have not done this in general. The extreme case is the simple sample with $n = 3$, $\nu = 2$, $\nu_0 = 0$. Rule 1 is now the same as Rule 0, because at most one observation in the sample can differ from the mean by more than Cs (when $C > 1$); and it follows that formulas (11.2) and (11.3) give exact results for Rule 1. For a premium of 2% we find we must take $C = 1.154638$, and then $\alpha = 0.00667$. The rejection rule can be expressed in the form: reject either extreme observation if its distance from the median observation exceeds 82 times the distance of the other extreme observation from the median. If one of the three observations has a bias of magnitude $b\sigma$, the other two being good, the chance that the rejection rule will cause the exclusion of the bad observation is about $\frac{1}{2}$ if $b = 79$; and b needs to be three or four times this size before rejection of the bad observation can be said to be virtually certain. For all practical purposes the rejection rule is utterly useless and absurd (cf. Lieblein, 1952). One may conjecture that a Studentized rejection rule will have low power whenever $\nu + \nu_0$ is small, say less than 30.

References

Bertrand, J. (1888). *Calcul des Probabilités*. Paris. §166.

Bessel, F. W. and Baeyer, J. J. (1838). *Gradmessung in Ostpreussen*. Berlin. (*Abhandlungen von F. W. Bessel*, vol. 3, Leipzig, 1876; quoted by Czuber and Wellisch.)

Bliss, C. I., Cochran, W. G. and Tukey, J. W. (1956). A rejection criterion based upon the range. *Biometrika*, 43, 418–22.

Brownlee, K. A., Kelly, B. K. and Loraine, P. K. (1948). Fractional replication arrangements for factorial experiments with factors at two levels. *Biometrika*, 35, 268–76.

Chauvenet, W. (1863). *Manual of Spherical and Practical Astronomy*. Philadelphia. Appendix on the method of least squares, §§ 57–60. (The appendix was reprinted as *A Treatise on the Method of Least Squares*, 1868.)

Czuber, E. (1891). *Theorie der Beobachtungsfehler*. Leipzig. First part, §11.

Edgeworth, F. Y. (1887). The choice of means. *Philosophical Magazine* (5), 24, 268–71.

Federighi, E. T. (1959). Extended tables of the percentage points of Student's t-distribution. *Journal of the American Statistical Association*, 54, 683–8.

Glaisher, J. W. L. (1873). On the rejection of discordant observations. *Monthly Notices of the Royal Astronomical Society*, 33, 391–402.

Grubbs, F. E. (1950). Sample criteria for testing outlying observations. *Annals of Mathematical Statistics*, 21, 27–58.

Jeffreys, H. (1939). *Theory of Probability*. Oxford. §§4.4, 4.41, 5.77, 7.2. (In the second edition, 1948, the old §5.77 is recast and numbered §5.7.)

Kruskal, W. H. (1960). Some remarks on wild observations. *Technometrics*, 2, 1–3.

Lieblein, J. (1952). Properties of certain statistics involving the closest pair in a sample of three observations. *Journal of Research of the National Bureau of Standards*, 48, 255–68.

Pearson, E. S. and Hartley, H. O. (eds.) (1954). *Biometrika Tables for Statisticians*, Vol. 1. Cambridge.

Peirce, B. (1852). Criterion for the rejection of doubtful observations. *Astronomical Journal*, 2, 161–3.

Plackett, R. L. and Burman, J. P. (1946). The design of optimum multifactorial experiments. *Biometrika*, 33, 305–25.

Rider, P. R. (1933). *Criteria for Rejection of Observations* (Washington University Studies, New Series, Science and Technology, No. 8). St. Louis.

"STUDENT" (1927). Errors of routine analysis. *Biometrika*, 19, 151–64.

THOMPSON, W. R. (1935). On a criterion for the rejection of observations and the distribution of the ratio of deviation to sample standard deviation. *Annals of Mathematical Statistics*, 6, 214–9.

TUKEY, J. W. (1949). One degree of freedom for nonadditivity. *Biometrics*, 5, 232–42.

TUKEY, J. W. (1960). A survey of sampling from contaminated distributions. *Contributions to Probability and Statistics: a Volume Dedicated to Harold Hotelling*. Stanford.

WELLISCH, S. (1909). *Theorie und Praxis der Ausgleichungsrechnung*. Wien. Vol. 1, §40.

WRIGHT, T. W. (1884). *A Treatise on the Adjustment of Observations by the Method of Least Squares*. New York. §§69–73.

Locating Outliers in Factorial Experiments

CUTHBERT DANIEL

New York City

Editor's Note: The following paper by Mr. Cuthbert Daniel was presented at the Washington, D. C. Meetings of the American Statistical Association on December 29, 1959, as part of a session on the subject of outliers. Mr. Daniel later became aware of an error (see footnote associated with Equation 6) that may invalidate some of the conclusions reached in the paper. Nevertheless, the Editor has decided to publish the paper so that a complete record of the meeting is available, and in the belief that the paper represents an important beginning in the study of patterns of residuals resulting from carefully plan experimental designs.

1. INTRODUCTION

An outlier in a factorial experiment is an observation whose value is not in the pattern of values produced by the rest of the data. The pattern has two aspects, and two kinds of disturbance of the pattern may appear. That observation is suspect whose removal greatly simplifies the description of the rest of the data. That observation is also suspect that complicates the description of the error distribution.

An outlier may be due to a single relatively large interaction in one cell; it may be an extreme manifestation of the random error system; or it may be a mistake. Unless they are quite large, unreplicated factorial experiments can not be counted on to supply positive evidence for deciding between these three alternative possibilities. Partially replicated factorial experiments permit an unbiased and independent error estimate, and so put the experimenter in a better position to spot bad values and to decide on their source.

An advantage of full replication (I believe not mentioned by Fisher) is the security against outliers given by having at least two independent observations for each treatment combination. The empirical cumulative distribution of the ranges of the replicate pairs is often all that is necessary to detect a bad value and to decide that a mistake has been made (Mistakes are in my experience the commonest source of bad values. Such mistakes are usually not clerical but technical.) When the error distribution is normal, the ranges of pairs have the half-normal distribution and so can be plotted about a straight line on half-normal probability paper (1). The magnitude of the largest range or two can then be put in proper perspective.

Bad values are, then, easily found in fully replicated experiments. A strong case for such replication can be made when there is prior evidence of frequent bad values. But this evidence is often not available and the expermienter may then decline to spend much of his time in replication. In my judgment the statistician should always urge at least partial duplication (2, 3), especially

in the smaller fractional replicates of the 2^{p-q} form when $(p - q)$ is 4 or 5. In larger experiments, and when the run-to-run variance is securely known, single replicates may be all that can be managed. This paper deals mainly with unreplicated experiments.

2. One Biased Value in a Two-Way Layout

The row and column classifications are here assumed discontinuous or qualitative. The first panel of Table 1 shows a set of (synthetic) data in a 5×4 layout. The second panel shows the "regression values" computed from the row averages and the column deviations under the assumption of additivity. The third panel gives the residuals, that is, the differences between the original values of the first panel and the corresponding regression values of the second.

The second row of panels in Table 1 shows the same set of data, handled in the same way, but with *one* value (third row, third column) decreased by 20. It is seen that the largest residual is in the cell (3, 3). However it is not 20 less than the former residual but only 12 less.

The third row of this table shows the effects of this one change in the 5×4 layout. The last row in Table 1 shows the effects on the residuals of a bias of size RC in a single cell (here chosen as (1, 1)) of an $R \times C$ layout. Here r is $R - 1$ and c is $C - 1$. It is obvious that the same pattern can be centered on any cell.

A biased value will produce a pattern in the calculated discrepancies, d_{ij}, that should show high correlation with the constants, c_{ij}, given at the end of Table 1, when the latter are centered on the maximum discrepance, $d_{\max}$. The numerator of the correlation coefficient,

$$r_{c_{ij}.d_{ij}} = \left(\sum_{ij} c_{ij}\, d_{ij}\right) \Big/ \left(\sum c_{ij}^2 \cdot \sum d_{ij}^2\right)^{1/2} \tag{1}$$

is, as always a contrast. We can re-write equation 1

$$r^2_{c_{ij}.d_{ij}} = \hat{\theta}_m^2 / (rcRC)(SSD) \tag{2}$$

where SSD is the sum of squares of the discrepances and

$$\hat{\theta}_m = \sum_{i=1}^{R} \sum_{j=1}^{C} c_{ij}\, d_{ij} = \sum_{i=1}^{R} \sum_{j=1}^{C} c_{ij} y_{ij} = RC\, d_{\max} \tag{3}$$

is the "maverick contrast". The third member of equation 3 is useful for estimating the variance of the maverick contrast because the observations, y_{ij}, are often safely assumed statistically independent. Thus $\mathrm{Var}\,(\hat{\theta}_m) = \sigma^2 \sum\sum c_{ij}^2$. The last member reveals that all the information about a single bad value is in the largest residual, $d_{\max}$, after all. Using the identity

$$\sum_{i}^{R} \sum_{j}^{C} c_{ij}^2 = rcRC, \tag{4}$$

the variance of the largest discrepance is seen to be rc/RC times the error variance.

Substituting the values given by equations 3 and 4 in equation 1, and abbreviating the mean square for all RC residuals by MSD, we get

$$(5) \qquad r^2_{e_{ij},d_{ij}} = \frac{RC}{rc} \cdot \frac{d^2_{\max}}{SSD} = \frac{RC}{r^2c^2} \frac{d^2_{\max}}{MSD}$$

which shows how r^2 varies with the size of the layout and with the observed discrepancies. This formula for r^2 shows the fractional reduction in the residual sum-of-squares, *SSD*, that is possible by revision of the response in the cell showing the largest residual. For the data of Table 1, as modified in the second set of panels, this fraction is 0.669.

TABLE 1

Effects of one biased value on residuals in an additive two-way layout

	"Data"				Avs.	Regression				Residuals			
	35	32	40	37	36	32	33	42	37	3	−1	−2	0
	29	29	36	34	32	28	29	38	33	1	0	−2	1
	25	29	40	30	31	27	28	37	32	−2	1	3	−2
	19	25	35	25	26	22	23	32	27	−3	2	3	−2
	22	20	29	29	25	21	22	31	26	1	−2	−2	3
Avs.	26	27	36	31	30								
Devs.	−4	−3	6	1									

	Biased Data				Avs.	Biased Regression				Biased Residuals			
	35	32	40	37	36	33	34	39	38	2	−2	1	−1
	29	29	36	34	32	29	30	35	34	0	−1	1	0
	25	29	*20*	30	26	23	24	29	28	2	5	−9	2
	19	25	35	25	26	23	24	29	28	−4	1	6	−3
	22	20	29	29	25	22	23	28	27	0	−3	1	2
Avs.	26	27	32	31	29								
Devs.	−3	−2	3	2									

	Bias in Data				Avs.	Bias of Regression				Bias of Residuals			
	0	0	0	0	0	1	1	−3	1	−1	−1	3	−1
	0	0	0	0	0	1	1	−3	1	−1	−1	3	−1
	0	0	−20	0	−5	−4	−4	−8	−4	4	4	−12	4
	0	0	0	0	0	1	1	−3	1	−1	−1	3	−1
	0	0	0	0	0	1	1	−3	1	−1	−1	3	−1
Avs.	0	0	−4	0	−1								
Devs.	1	1	−3	1									

	RC	0	0	...	*R*	*rc*	−*r*	−*r*	−*r*	·	·
	0				0	−*c*	1	1	1	·	·
	0				0	−*c*	1	1	1	·	·
	·				0	−*c*	1	1	1	·	·
	·				0	·	·	·	·	·	·
	·				0	·	·	·	·	·	·
Avs.	*C*	0	0	···	1						
Devs.	*c*	−1	−1	···							

If we had only the spoiled data, and did not know that 20 had been removed from the observation at (3, 3), we would have spotted its location from the large discrepance. We would then have estimated a simplifying value to enter in its place. The easiest way to do this is to revise the response by the amount $-RC/rc$ times the discrepance d_{max} . For the numbers given the change would be plus 15. Thus the present 20 would be changed to 35. The new SSD is $(1 - r^2)$ times the old SSD. For our case the new SSD is then $(1 - 0.669)202$ or 67. The new estimate of sigma, called s_1 , is $(67/11)^{\frac{1}{2}}$ or 2.47. The earlier estimate, based on the biased residuals would be $(202/12)^{\frac{1}{2}}$ or 4.10.

Since the contrast $\hat{\theta}_m$ is chosen after the data are taken, it seems natural to use a modification of Scheffé's S-method (pages 68–72, reference 4) for judging the reality of the bias, and hence of the largest residual. It will be assumed in this section that a fair estimate of the true error variance can be obtained from the revised discrepances, with $(rc - 1)$ degrees of freedom. This estimate, called s_1^2 below, can be computed from Equation (6)[1].

$$s_1^2 = \left(SSD - \frac{RC}{rc} d_{max}^2\right) \Big/ (rc - 1) \qquad \text{(see footnote)} \tag{6}$$

Scheffé's method permits us to find critical magnitudes of the maverick contrast, that is, values that will be judged just significant at some chosen confidence level. Using primes on $\hat{\theta}_m$ and on d_{max} to denote just-significant values, we place

$$\theta_m' = S\hat{\sigma}(\hat{\theta}_m); \tag{7}$$

where

$$S = (rcF_0)^{\frac{1}{2}}, \tag{8}$$

and F_0 is the $(1 - \alpha)\%$ critical value of F, that is

$$F_0 = F_{(1-\alpha),rc,rc-1} \; ; \tag{9}$$

and where the estimated standard error of the maverick contrast is given by

$$\hat{\sigma}(\hat{\theta}_m) = s_1(rcRC)^{\frac{1}{2}}. \tag{10}$$

Substituting the values given by equations 4, 8, 9 and 10 in equation (7) gives the following simple relationship between d_{max}' and s_1 :

$$RC\, d_{max}' = s_1(rcF_0)^{\frac{1}{2}}(rcRC)^{\frac{1}{2}}. \tag{11}$$

Defining $t_{1,\alpha}$ as the ratio of these two quantities as shown in equation 12,

$$t_{1,\alpha} = d_{max}'/s_1 = rc(F_0/RC)^{\frac{1}{2}}, \tag{12}$$

[1] H. Scheffé and H. A. David have pointed out a serious error here. If there is an outlier at position (i, j) then the assumption seems fair that s_1^2/σ^2 is distributed as chi-square with $(rc - 1)$ degrees of freedom. If however there is *no* outlier then the mean-square of all rc residuals (MSD) is distributed as chi-square with rc degrees of freedom, and s_1^2 is *not* chi-square distributed with $(rc - 1)$ degrees of freedom. I do not now know how great are the effects of this error on the values given in Tables 1 and 2. The Editor has asked that the paper be published as given in the interests of having all papers and discussion in one issue. Revised tables will be published in this Journal as soon as possible.

we get the critical values given by the third member of that equation to compare with the statistic t_1, the corresponding quantity obtained from data, and defined by equation 13.

$$t_1 = d_{max}/s_1 . \tag{13}$$

Table 2 gives some approximate critical values of this ratio for an α *per experiment* of 0.2.

TABLE 2

Values of $t_{1,\alpha}$ ($\alpha = 0.2$) *for* $R \times C$ *layouts.* R *and* C, 3 *to* 9.

R \ C	3	4	5	7	9
3	2.29	2.58	2.87	3.40	3.84
4		3.06	3.48	4.18	4.78
5			3.80	4.83	5.36
7				5.93	6.89
9					7.88

The statistic t_1 found from the data of Table 1 is 9/2.47 or 3.65, which is greater than the 3.48 given in Table 2. With a 20% risk of false alarm, then, the value at (3, 3) would be judged an outlier.

3. One Biased Value in a 2^n Unreplicated Experiment

A further complication arises here in that the regression equation is to be derived from the data and is not specified beforehand as it is in the two-way layout.

Routine analysis of the unrevised data may not suffice to give clues to the real effects. This will generally be the case if a defective value which has an expected value near one end of the range of responses actually appears near the other end of the range. I see no alternative in such cases to making a series of trials, replacing in turn each of the 2^n responses by the average of the others, and then revising the estimates of all effects accordingly. If one and only one of the resultant sets of effects permits a simple regression equation, containing mostly main effects and interactions among the factors having large effects, then a strong case can be made for naming as defective the value whose revision caused the simplification. If none of the substitutions proposed produces a notable simplification, two alternatives suggest themselves. There may be two bad values, which can be searched for using clues given by the one-at-a-time analysis just described. This failing, a half or full replication of the experiment may be necessary. But these grim contingencies are relatively rare. More common is the appearance of a single bad value, large enough to cause trouble, but not so large or so placed as to conceal all real effects.

If a bias of 16 were to hit the treatment combination *abcd* in a 2^4 experiment, the expected pattern of residuals, regressing only on the mean, is given by the first column of numbers in Table 3. If a different observation, not *abcd*, were

TABLE 3

Effects on residuals of a bias of 16 at abcd in a 2^4 exp.

	Successive regression equations						
	μ	$\hat{A}$	$\hat{B}$	$\hat{C}$	$\widehat{AB}$	$\hat{D}$	$\widehat{ABCD}$
(1)	−1	0	1	2	1	2	1
a	−1	−2	−1	0	1	2	3
b	−1	0	−1	0	1	2	3
ab	−1	−2	−3	−2	−3	−2	−3
c	−1	0	1	0	−1	0	1
ac	−1	−2	−1	−2	−1	0	−1
bc	−1	0	−1	−2	−1	0	−1
abc	−1	−2	−3	−4	−5	−4	−3
d	−1	0	1	2	1	0	1
ad	−1	−2	−1	0	1	0	−1
bd	−1	0	−1	0	1	0	−1
abd	−1	−2	−3	−2	−3	−4	−3
cd	−1	0	1	0	−1	−2	−3
acd	−1	−2	−1	−2	−1	−2	−1
bcd	−1	0	−1	−2	−1	−2	−1
*abcd	15	14	13	12	11	10	9
S. S.	240	224	208	192	176	160	144
M. S.	16	16	16	16	16	16	16

too large by 16, the residual 15 would of course shift to the corresponding row, all other rows showing −1. If, still assuming one value too large by 16 at *abcd*, we had found that effect A was large, we would have computed discrepances from a regression equation that contained an estimated mean and an estimated A-effect. The expected values of the resulting biased residuals are given in Table 3 in the column headed $\hat{A}$. If the mean, and A and B were judged real, the discrepances would have the biases appearing in the following column. Similar remarks hold for the sequence of regression equations obtained by successively including the terms at the heads of the columns.

The magnitude of the largest bias decreases with $(16 - K)$, where K is the number of terms in the regression equation. Thus it might be thought that the inclusion of more terms will give a better fit to data that contains one bad value. However, the mean square for lack of fit does *not* continue to decrease, but settles down to a nearly constant value as indicated by the limiting case exhibited in Table 3, where the random errors are all zero. This persistent failure of the successive equations to fit better may well be a good indicator of the presence of a bad value. It would *not* be an assurance that our mean square was converging on the true variance.

In a 2^n experiment, a bad value enters into every contrast, either increasing or decreasing its value by the amount of its bias. The average absolute value

of the contrasts will be increased, and so a half-normal plot (1) of the ordered contrasts will have too few contrasts near zero. A straight line through most of these contrasts does not go through the origin as it should but hits the abscissa axis at a positive value. This graph can be used, then, even before any discrepances are plotted, to warn us of a single biased value. It has the advantage that the detection does not depend on which effects are judged to be real.

The pattern of biases depends on the location of the biased value and on the regression equation chosen to represent the data. The pattern can be quickly written down for any combination of interest with the aid of the standard transformation matrix of plus and minus ones that specifies how each observation enters each effect-contrast. Such a matrix is given in Table *M* of the book, *Design and Analysis of Industrial Experiments*, edited by O. L. Davies. It is only necessary to mark the line corresponding to the suspected value and the columns corresponding to the desired regression equation. After reversing the signs of all entries in the columns in which the bias enters negatively, it is only required to add the plus and minus ones in each row to get the expected values of the propagated bias. Subtracting these values from the actual bias (a column of zeroes and one 16) gives the biases in the discrepances which are the c_i listed in Table 3 and used in equations 14 and 15 below.

Exactly analogous to the maverick contrast for two-way layouts, one can be written down for the 2^N experiments. Thus

$$\hat{\theta}_m = \sum_{i=1}^{2^N} c_i\, d_i = 2^N d_{\max} \tag{14}$$

The variance of this contrast is

$$V(\hat{\theta}_m) = \sigma^2 \sum_{i=1}^{2^N} c_i^2 = 2^N(2^N - k)\sigma^2 \tag{15}$$

By reasoning parallel to that outlined above in equations 7 through 13, the critical ratio

$$t_{1,\alpha} = (2^{N-k}/2^{N/2})F_0^{\frac{1}{2}} \tag{16}$$

can be derived. Table 4 gives some approximate values for this ratio for experiments with 16, 32 and 64 run factors and with 2, 4, 6 or 8-term regression equations.[2]

TABLE 4

Values of $t_{1,\alpha}$ ($\alpha = 0.2$) for 2^n experiments; K terms in regression

K \ n	4	5	6
2	4.42	6.16	8.45
4	3.89	5.82	8.18
6	3.34	5.42	7.94
8	2.78	5.07	7.70

[2] See previous footnote.

References

[1] Daniel, C., Use of Half-Normal Plots in Interpreting Factorial Two-Level Experiments, *Technometrics* 1, No. 4.
[2] Daniel, C., Fractional Replication in Industrial Experimentation, Proceedings National Convention, American Society for Quality Control, 1957.
[3] Dykstra, O., Partial Duplication, *Technometrics* 1, 63–75.
[4] Scheffé, H., *The Analysis of Variance*, Wiley and Sons, New York, 1959.

Discussion of the Papers of Messrs. Anscombe and Daniel

Editor's note: The two preceeding papers in this issue of *Technometrics:* "Rejection of Outliers" by F. J. Anscombe and "Locating Outliers in Factorial Experiments" by C. Daniel, were presented at the Washington D. C. meetings of the American Statistical Association on December 29, 1959 at a session on the general topic of "Outliers." The session was part of the program organized by the Section of the Physical and Engineering Sciences of the ASA. The papers were discussed by W. Kruskal, T. S. Ferguson and J. W. Tukey. At the conclusion of the meetings the Editor proposed that the papers and their discussion be published in *Technometrics* and all participants agreed. Professor E. J. Gumbel, who was unable to speak at the meetings because of limitations of time, was also invited to present his comments. In a subsequent study, Mr. Daniel discovered an error in his paper that had been missed by the reviewers. The error is noted in Mr. Daniels paper in this issue of *Technometrics.*

WILLIAM KRUSKAL

University of Chicago

The two papers of this session are obviously outliers among the population of papers appearing at statistical meetings. Further, these outliers are clearly in the *positive* direction. I have established this by use of a subjective rule, which considers how much papers stimulate me to think about the problems they treat and how much new ways of thinking about the problems are suggested.

Both of the papers depart from the conventional treatment of outliers in that more complex situations than simple random normal samples are contemplated. Specifically, both papers discuss outliers primarily in the context of Model I analysis of variance designs with a high degree of symmetry. Both papers deal with rules for rejecting apparent outliers. The two papers differ in that Daniel's is mainly concerned with rejection rate and power, while Anscombe's is mainly concerned with the mean square error of estimators following application of a rejection rule. I think that both papers help towards closing the ubiquitous gap between practice and theory.

As to practice, I suggest that it is of great importance to preach the doctrine that apparent outliers should *always* be reported, even when one feels that their causes are known or when one rejects them for whatever good rule or reason. The immediate pressures of practical statistical analysis are almost uniformly in the direction of suppressing announcement of observations that do not fit the pattern; we must maintain a strong sea-wall against these pressures. The extra work and publication space needed to describe apparent out-

liers and the reasoning behind their treatment is usually negligibly small.

Second, still about practice, one of the most important results of finding an outlier can be the correction of a flaw in the measurement process, or—even better—the creation of new insights into the phenomena under study. For such purposes, it may be that something like significance testing is more germane than evaluation of mean square error.

In this general connection, I should like to call attention to a debate between H. C. Hamaker, on the one hand, and R. C. Anderson and A. W. Wortham, on the other. This appeared in the July, 1959, issue of *Industrial Quality Control.*

I have two comments about theory. First, there may be cases in which the probability of a blunder *depends on* the value that would have been observed if the blunder were not present. For example, a scale reading may be more subject to gross misreading if it is towards the end of the scale than if it is near the middle. Can reasonably realistic models for this be stated and worked on? This question is closely related to another that I think deserves more attention: suppose that the probability that an observation is *missing* depends upon the value that would have been observed; to what extent are we justified in applying traditional missing-value manipulations?

My final comment is more technical. In both papers, rejection rules are contemplated that sometimes lead to analyses with high symmetry *except for* one or two missing observations, missing now because of rejection by some rule. We know how to handle such situations with relative ease by computing the least-squares estimate of the missing observation's expectation, putting the result in the sample in lieu of the missing observation, and then proceeding with estimation as if the full sample were there. These all-or-nothing rejection rules may have been proposed because of the ease of analysis following rejection of one or two apparent outliers.

I suggest that alternative rules are worth consideration, in which the apparent outlier is not banished from the company of decent observations, but rather is given a lower weight. This can be done with hardly any more trouble than finding the so-called replacement for the missing value that one would do after rejection.

Specifically, suppose that the sample is $y_1, \cdots, y_n$, with mean specified to lie on a given linear manifold in n-space. Suppose, to begin with, that $y_1, \cdots, y_{n-1}$ all have variance 1, but that y_n has variance σ^2, a known number. (It would be enough if only the ratio between σ^2 and the common variance of $y_1, \cdots, y_{n-1}$ were known.) Let y be the estimate of the expected value of y_n that one obtains following the usual missing-value approach for y_n. Then one may state a linear combination of y and y_n, $(y + \lambda y_n)/(1 + \lambda)$, such that, if you replace y_n by this linear combination and apply traditional analysis to the altered sample, you get the appropriate least-squares estimators with proper weighting for y_n. The weight λ is readily obtained in terms of the quantities needed to find y.

This may be derived algebraically by the use of M. A. Woodbury's result on inverting special matrices, or it may be obtained geometrically. Extensions to more than one separately weighted y_i are undoubtedly possible in various ways, and I hope to return to this topic in more detail elsewhere.

THOMAS S. FERGUSON

U. C. L. A. and Princeton University

I wish to thank and compliment the speakers for producing first rate work on the theory of rejection of outliers. Both papers treat a topic which has not even been attempted before in the outlier literature, namely, the problem of rejection of outliers in designed experiments, where the means of the observations may be complicated functions of unknown parameters so that even discovery of a bad value may be difficult. The methods of the two papers differ. Whereas Daniel makes use of "patterns" which appear in the residuals, Anscombe goes straight to the extreme residual for his rejection statistic. The fact that the correlation test that Daniel proposes turns out to be equivalent to the studentized extreme residual is both surprising and gratifying in that it is a studentized justification of the rejection rule of Anscombe, at least in the two designs considered. One suspects that this could be extended at least to all designs considered by Anscombe. Another justification can probably be made by extending the result of Paulson ("An optimum solution to the k-sample slippage problem for the normal distribution," *Annals of Math. Stat.* 1952,) to two-sided rules and for designed experiments.

Daniel's treatment of the two-way layout is a nice neat package. I want only to say that I don't think one can avoid using a multiple comparison method here. The extreme residual is not a simple contrast, but, however, it is a function of all the simple contrasts.

Finding outliers in factorial experiments is inherently more difficult because one doesn't know which effects are going to have large mean values. Furthermore, one maverick may mask the real effects so that we may not be able to judge which effects to take into the regression equation. The use of the half-normal plot here is ingenious. However, one needs a more satisfactory way of finding which of the original observations is the bad one. The original half-normal plot may tell you that there is a spurious observation but one would hate in a 2^5 experiment to have to make 32 more plots, one each for treating each original observation in turn as a missing value, in order to determine which observation is the bad one.

A related point is of interest here: that of chosing the plotting points for the half-normal plot. Suppose we have an ordered sample of size n from a half-normal distribution $x_1 < x_2 < \cdots < x_n$. For the half-normal plot we plot the points (x_i , y_i) in the plane, where the y_i are chosen beforehand so that the points will in some sense lie on a line. It is usual to take y_i as that number such that $F(y_i) = i/(n + 1)$ (or $F(y_i) = (i - \frac{1}{2})/n$) where F is the distribution function of the half-normal distribution. I would prefer to use y_i either as the expected value of x_i , or as the ith entry of the eigenvector corresponding to the largest eigenvalue of the second moment matrix of the x_i . The first has the property that x-intercept of the least squares regression line of x on y is an unbiased estimate of zero. The second has in addition to this the property that the expected value of the residual sum of squares of the distances to this least squares regression line is a minimum. However, tables of the expected values and of the second moments of the order statistics for the half-normal

distribution do not exist. I would like to appeal to the statistical public to compute such tables.

I feel that Anscombe's paper will be the cornerstone on which future generations of outlier rejectors will base their work. The basic idea, which appears is all statistics, is that one can better design the experiment if one knows what actions one will take later. This is a point which decision theory emphasizes. However the proferred solution to this problem took an extra bit of ingenuity; it is not decision theory but insurance. I would like to make only one remark, that Rule 0, studentized, is exactly Thompson's rule, (*Annals of Math. Stat.* 1935) and does not suffer from the defect mentioned in the paper for the nonstudentized. I think it is a serious contender for attention in the extension of these ideas to the problem when the variance is unknown.

JOHN W. TUKEY*

Princeton University and Bell Telephone Laboratories, Murray Hill

I would like to begin by saying that I agree with all the good things that the earlier discussants have said about the papers.

This morning's meeting illustrates how important it is for us to return to the "Elementary Problems of Statistics". One of the previous discussants suggested that the two papers would be foundation stones for the future work on the rejections of outliers. Notice that these papers were not thought of as finishing a process but rather as a beginning, as initial steps in further growth of work on outliers. Surely the rejection of outliers is one of the really elementary problems of statistics. But, as we are coming to recognize, it does not represent a problem which is anywhere nearly completely treated.

In connection with the discussion of whether a rejection procedure should be a test of significance or not, I think that it is important to emphasize the way in which Frank Anscombe's paper avoided the semantics of "rejection". He was concerned with doing the best that he could with the data before him, with combining it in the best way that seemed reasonably feasible ... not with whether an individual observation should be dropped or should not be dropped, not with whether an individual observation was in some specific sense unusual or not ... only with the question of what was the best thing to do. By asking how well one does in estimation, one avoids any implication of making up one's mind whether or not a particular observation is a "bad value", one avoids any need for "significance testing".

Bill Kruskal has pointed out the connection between such yes-or-no procedures as rejection and the much more flexible and quantitative procedures of weighting. (At this point it may be appropriate to notice one further related reference, a paper by H. R. Hulme and L. S. P. Symms called "The Law of Error and the Combination of Observations", Monthly Notices of the Royal Astronomical Society, June 1939, Vol. 99, pp. 642–649. This paper discussed the use of a weighting procedure in reducing some 9,000 observations made with a floating zenith telescope at Greenwich over a ten-year period. It is, of course, related

* Prepared in part in connection with research at Princeton University under Contract No. DA 36-034-ORD-2297 sponsored by the Office of Ordnance Research.

to the Jeffreys techniques referred to in Frank Anscombe's paper.) There are *three qualitative operations* which can be used to deal with outliers and wild values. It may well be that the best practical solutions will turn out to involve combinations of these. These are (i) rejection, (ii) truncation, and (iii) Winsorization. We have heard a lot today about rejection, where one chooses which extreme values to delete on the basis of the specific and quantitative relation of these values to the other values observed. In symmetrical truncation, by contrast, one deletes a certain percentage of high values and an equal percentage of low values, without regard to just how the values compare with one another.

Finally there is Winsorization. Charles Winsor put forward a principle of quite general application, namely: While the numerical value of an apparently wild observation is untrustworthy, the direction of its deviation (e.g. high or low) is worthy of some attention. He applied this principle to outliers by taking the largest deviations or largest residuals and decreasing their magnitude, while retaining their sign, until they are equal to the next largest ones, thus making a qualitatively reasonable adjustment. In fact, it is possible to give quite quantitative reasons why this sort of adjustment is not only a convenient approach but an effective one.

The natural way to attack the class of problems of which defense against wild values is a part, is not to begin in small samples with rejection, but to begin in large samples with Winsorization and truncation. However one may, and I believe should, look forward to an eventual tendency to use a combination of two or more of these procedures in all situations. The tentative suggestion in Cuthbert Daniel's paper that, for example, in a 2^5 factorial experiment one might plan to reject one observation all the time and a second one some of the time, is precisely a proposal to combine some truncation with some rejection.

(I would like to say that I agree with him both in putting forward this idea and very particularly, in putting it forward tentatively. It seems to me this is the sort of procedure that must be considered, but it is also clear that it is the sort of procedure which, under some circumstances, might give rise to difficulty. We will have to think about this sort of procedure carefully, and try it out in practice, before we know when and where it can be routinely adopted.)

I would like now to say a few things about the case where σ is known. The results in Anscombe's Table 3 can be put into somewhat more compact form if one notices that, for fixed premium and to an adequate approximation, C is proportional to $(\nu/n)^{0.4}$, α is proportional to $(\nu/n)^{-1.2}$, and b is proportional to $(\nu/n)^{0.6}$. If we take advantage of such facts we can greatly compress the tables upon which we are to base our judgments.

Frank Anscombe emphasized the size of the premium that had to be paid, and he also emphasized that the amount of protection provided had also to be considered. I suspect that the protection here is perhaps best thought of in terms analogous to the collison insurance of automobiles, in terms of "what is the amount deductible; what is the largest net loss which one may have to suffer?" It would clearly be of interest to know just how much of an increase in variance one is going to suffer in the worst situation.

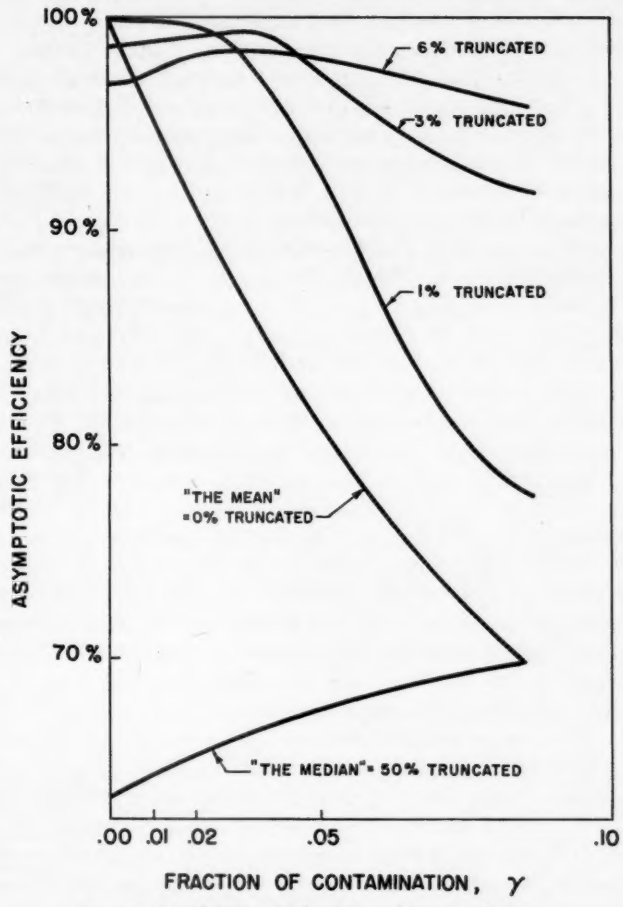
It is important to realize that there are many different sorts of difficulty.

The particular sort of difficulty that has been emphasized in Anscombe's paper is exactly that which arises when a normal distribution is supplemented by two isolated masses of probability, one at $+b\sigma$ and one at $-b\sigma$, and these isolated masses are just big enough so that one observation drawn from them will occur in at least a moderate fraction of samples of the chosen size, but two observations per sample will occur quite infrequently. Clearly one must think about different values of b. It even may be that one can think about a probability distribution for these values of b. But it is doubtful that *any one* probability distribution would suffice. Throughout the discussion of outliers and wild values we have got to face a *variety*, all tentative, of bad possibilities.

As some of you in the audience know, I tend to consider cases where the badness is represented by the mixture of two normal distributions, with the same mean but with different variances (where we have a contaminated distribution), much more than the case where we have a normal distribution plus some specific values with finite probability. Here, too, one must consider different possibilities, exemplified by different degrees of contamination. Let us look at the case where we have a large sample and are worried about this other sort of difficulty. If we consider contaminated distributions in which the mixture includes a certain fraction γ of a wider normal distribution with three times the scale of the narrower normal distribution (which appears at a fraction $1 - \gamma$), then we have the results shown in the figure (taken from "A Survey of Sampling from Contaminated Distributions", Chapter 39 of *Contributions to Probability and Statistics*; *A volume dedicated to Harold Hotelling*, Stanford University Press, 1960) for various sorts of truncated means.

The curve labeled 0%-truncated is of course the curve for the ordinary mean, where nothing has been deleted from the observations. At the bottom we have the curve where we have removed the 50% highest observations and the 50% lowest observations, leaving only a single observation in the middle, so that the mean of the retained values turns out to be mean of a single central value, that is the median. The other curves show the result of cutting one, three, or six percent, respectively, of the observations off the upper end of the sample and an equal amount off the lower end of the sample. These results are expressed in terms of asymptotic variance and large sample theory; they are surely not applicable down to samples as small as say 10, though they may be fairly effective for samples of the order of 100. Surely, they do not work down to the level of samples of three or four, which Frank Anscombe pointed out are of frequent occurrence. Here the premium (in Frank Anscombe's terms) is the amount by which the curve lies below 100% for $\gamma = .00$. For as much as six percent truncation on each tail, the premium is only about 3%. For a 3% premium we can protect ourselves against a possible loss of about 30%. This example makes clear the possibility of using truncation in large samples.

Next a few words about the case where σ is unknown, the case where we are concerned either with regresson situations regarded as such, or with patterned experiment situations which can be treated from a regression viewpoint. While Cuthbert Daniel was clearly aware that there was no need to calculate the standard deviation of the cleaned-up observations in order to get something to compare with the maximum residual, I should like to emphasize that this



Asymptotic efficiency, for location, of truncated means.
(Figure reprinted with permission from *Contributions to Probability and Statistics*, Olkin et al., Copyright 1960, Stanford University Press.)

is the case. I should also like to emphasize that one can always compare the maximum residual with a suitable multiple of either the square root of the sum of squares of residuals, or with a different suitable multiple of the square root of the mean square for residuals. The actual numerical value of either suitable multiple will depend, in a two-way case, for example, upon the number of rows and columns used. (Such dependence is going to occur for the critical value of any ratio one chooses in this situation.) One can always compare the maximum residual with something calculated from all the residuals together.

I cannot refrain from joining the authors and earlier discussants in emphasizing again that all the most natural ways of dealing with the residuals tend to be mathematically equivalent; that using the correlation coefficient, or the relatively biggest residual, leads to the same basis for making choices whether or not to

reject. And I also cannot refrain from underlining the fact that, in the long run, we are going to need to know somewhat more about the joint sampling behavior, in a wide variety of situations, of residuals.

Comments, both by speakers and discussants, make it, I think, also necessary to say a word or two about some multiple comparisons questions. It has been suggested that it is necessary to use Scheffé's approach. It seems to me that it is only necessary to make some sort of allowance for the fact that, when we look at the largest residual, we are looking at the largest of a certain number of possibilities (which have a certain interrelation), that we are doing some selection. If one were to use the distribution of this largest residual itself to generate a multiple comparisons procedure, one would get a somewhat more stringent "5% level" than if one used an F-statistic-based multiple comparison procedure at a "5% level". And the latter, again, would be more stringent than a procedure in which, when there were R rows and C columns, we would say that we have RC residuals, so we will take a critical value of $|t|$ which corresponds not to 5% but to $(5/RC)\%$. There will be some differences between these three approaches, but the differences will be minor and unimportant. The important thing is that all allow for the problem of selection. Neither the detailed approach nor the precision of its application will be important.

There was a suggestion by the discussants that Cuthbert Daniel's paper was concerned with model 1 anova (model 1 analysis of variance). I have to point out that this is not really true, especially because one of the things with which he was explicitly concerned in his talk was the fact that the particular sort of regression, the particular sort of separation between accepted effects and residuals, which would be appropriate for a given set of data, would depend upon what the data looked like after one (or more) of the values had been deleted. To have flexibility in the way in which the data are going to be interpreted, flexibility which is directly associated with the data values themselves is not to have model 1 anova. It is to have something else, something which is harder to analyze from the point of view of outliers, something which is much more reasonable from the point of view of actual data analysis. One of these days, perhaps, it will have a specific name, but for the present it should suffice to point out that it is *not* model 1.

There are a few other specific points that might perhaps be mentioned. One is that, if (i) the experiments are performed one after another and (ii) blocking is not too serious a problem, then, rather than planning to reject the experiment with the largest residual and analyze the others, one might plan to keep a continuing check on how the results are coming out, so as to know in advance of the completion of all the experimental runs (i.e. experimental trials, or plots), which run or runs seem likely to give the largest residuals so as to be able to set aside these runs providing large residuals, and to repeat each of the corresponding experimental conditions twice as many times as before. This would give good basis for deciding whether such wild shots occur randomly, or whether they are really associated with certain combinations of treatments. This is basically one of the most important problems in connection with outliers, one which has already been stressed by Bill Kruskal. It is often very hard to determine whether one wants to set the apparent wild shot aside and produce a

good analysis of what is left, or whether, since such understanding will be more valuable than any possible increase in the accuracy of the pre-chosen estimate or estimates, one wants to cling to the wild shot above all else as something whose occurrence must be understood.

If, on the other hand, the experiments are performed in parallel time-wise, so that duplication of the *post facto* unusual ones is out of the question and only deletion is possible, this problem: "Which is the baby, and which is the bath-water?" becomes, of course, very much more stringent. In either circumstances, whenever the wild shots (those things which appear from a factorial point to be "wild shots") are such as to indicate that some other way of looking at the data analysis would be better, one must be prepared to give up the factorial approach.

At one time Cuthbert Daniel objected to the use of a log-log type of half-normal plot because it made the smallest effects too noticeable. Now he is concerned with values moderately far down in the size ordering of effects. Now that such values are used in judging whether or not there appears to be a wild shot present, there will perhaps be more interest in the log-log form of the half-normal plot.

Finally, in closing, I would like to ask a question of interest, but probably not immediately of serious import. When are we going to get a combination of the problems we discussed at the Pittsburgh meeting and the problems we discussed here? When are we going to have the first discussion of the rejection of outliers in *super*saturated patterns where the value of ν/n is already negative?

E. J. GUMBEL
Columbia University

The acceptance or rejection of outliers is an important decision problem for all observations. Thus, this question arises for any statistical distribution used in the analysis of numerical data.

The outliers are values which seem either too large or too small as compared to the rest of the observations. Thus, they are extremes. The difficulty of the decision arises from the fact that the absolute size of the extremes increases in general with the sample size. A value which may be rejected for a certain number of observations has to be accepted for a larger sample since the gap between the outlier and the rest will be filled by new observations, one or several of which may even exceed the old outlier. Therefore, any criterion for such a decision must be based on the sample size. Anscombe proposes, for a normal population with known σ, to reject the absolute largest among n observations x_n if

$$|x_n - \mu| > C\sigma,$$

where μ is estimated as the mean of the retained observations. If σ is unknown it is to be replaced by s.

Anscombe's notation does not bring out the fundamental fact that C increases with the sample size n. Therefore it should be written C_n . A table of C_n as a function of n for chosen values of α would be desirable.

Anscombe claims that for $n = 3$ and 4 observations assumed to be normal with a known standard deviation "this method is probably the best that can be

suggested." The validity of this statement depends of course on what we define as being the best. Logical consistency, numerical simplicity of procedure and other criteria may lead to different definitions for such an optimum.

All rejection criteria are based on the extreme values as linked to the sample size. The simplest procedure for the normal distribution seems to me to start from the distribution of the extreme normal values for n observations and to choose an appropriate probability α for the rejection. The choice of this α is arbitrary and the conventional value of .05 may be appropriate. This method can be generalized for any known initial distribution. If it is not symmetrical the two distributions of the largest and of the smallest values as functions of the sample size have to be considered separately. Then Anscombe's rules could be applied in a modified way to these distributions of the extremes.

The rejection of outliers on a purely statistical basis is and remains a dangerous procedure. Its very existence may be a proof that the underlying population is, in reality, not what it was assumed to be. The alleged homogeneity may not be true and the outlier may be the most important observation because it throws light on the way the observations were made.

In the history of this problem, the important contribution of P. R. Rider, "Criteria for rejection of observations", Washington University Studies No. 8,1933 and the literature given therein should be mentioned.

AUTHOR'S REPLY TO DISCUSSION

F. J. ANSCOMBE

This has been a valuable discussion, because a number of suggestions have been made that will be really helpful in further research, and also because it serves to set the two papers in a just perspective. There are so many matters that can be discussed regarding outliers, so many different situations in which a problem of outliers arises, that one must be careful to note just what topic or situation is under consideration at any point. The topics that Cuthbert Daniel and I have discussed are narrow and special, relative to the whole field of outlier debate—though they may have seemed broad enough while we worked on them.

In comment on Cuthbert Daniel's paper, I should like to express the hope that his idea of detecting outliers by examining half-normal plots will be pursued further. This appears to be an essentially different device from those based on the largest residual.

I hope that further work concerning outliers and residuals will be presented at the Fourth Berkeley Symposium.

In revising my paper for publication, I have made various small amendments, prompted by the above public discussion, by private discussion with Churchill Eisenhart, and by detailed and searching comments from Leonard J. Savage and William Kruskal. To all I offer my grateful thanks. Irwin Guttman and I regret that the program of computations undertaken last year was not quite completed. We hope to add a supplement to Table 4 in a later issue of this journal.

Tests for the Validity of the Assumption that the Underlying Distribution of Life is Exponential: Part II

BENJAMIN EPSTEIN

Stanford University and Wayne State University

Part I of this paper was published in the previous issue of *Technometrics*, (Vol. 2, No. 1, February 1960). In Part I Dr. Epstein describes several graphical and analytical procedures for testing the assumption that the underlying distribution of life is exponential. This portion of the paper contains numerical examples illustrative of these procedures.

INTRODUCTION

In Part I of this paper the following tests, useful to determine the validity of the assumption that the underlying distribution of life is exponential, are described:

1. A Graphical Procedure
2. The χ^2 Test of Goodness of Fit
3. A Criterion Based on the Conditional Distribution of Total Lives
4. A Test for Abnormally Early Failures
5. A Test for an Abnormally Long First Failure
6. A Test whether the Mean Life (or Rate of Failure) in the First Half of a Life Test Differs Significantly from the Mean Life (or Rate of Failure) in the Second Half of a Life Test
7. A Test for Whether or Not the Mean Life Fluctuates During the Life Test
8. A Special Procedure for Procedure (7) When $r = 1$
9. A Test Based on the Maximum F Distribution
10. Tests for Abnormally Long Periods in which there is No Failure
11. A Graphical Procedure Based on the Kolmogorov-Smirnov Test
12. A Test Based on Conditional Rate of Failure

The description of these tests is followed by general comments, a short summary, an appendix in which some of the more mathematical arguments are found, and finally by a bibliography. Part II of this paper provides numerical examples descriptive of the procedures given in Part I.

EXAMPLES

(1) 49 items were placed on life test and the test was run until all items failed. The observed failure times $\tau_1, \cdots, \tau_{49}$ were:

1.2	13.7	38.9	72.4	102.8	151.6	203.0
2.2	15.1	47.9	73.6	108.5	152.6	204.3
4.9	15.2	48.4	76.8	128.7	164.2	229.5
5.0	23.9	49.3	83.8	133.6	166.8	253.1
6.8	24.3	53.2	95.1	144.1	178.6	304.1
7.0	25.1	55.6	97.9	147.6	185.2	341.7
12.1	35.8	62.7	99.6	150.6	187.1	354.4

Test the hypothesis that the underlying distribution of life is exponential, using procedure 1.

Solution: According to procedure 1, the graph of τ_i against

$$y_i = \log\left(\frac{n+1}{n+1-i}\right).$$

should be approximately a straight line passing through the origin, if the underlying distribution is exponential.

In our example we plot τ_i against $(n + 1)/(n + 1 - i)$ on semi-log paper (see Figure 1).

(2) 49 items were placed on life test and the test was run until all items failed. The observed failure times $\tau_1, \cdots, \tau_{49}$ were:

162	287	539	884	1178	1666	2181
172	301	629	886	1205	1676	2193
199	302	634	918	1377	1742	2445
200	389	643	1008	1406	1818	2741
218	393	682	1101	1501	1896	3191
220	401	706	1129	1576	2002	3577
271	508	777	1146	1641	2021	3694

One suspects that the waiting time for the first failure was abnormally long, and that the underlying p.d.f. is $1/\theta\, e^{-(x-A)/\theta}$, with $A > 0$. Give a graphical procedure for testing whether $A > 0$.

Solution: Again, as in the previous example, we plot τ_i against

$$y_i = \log\left(\frac{n+1}{n+1-i}\right).$$

As before, the data are plotted on semi-log paper (Figure 2). If the underlying distribution is one-parameter exponential, we should get a straight line passing through the origin (0, 0); if the underlying distribution is two-parameter exponential, we should get a straight line passing through $(A, 0)$. In our example it is fairly evident that we would reject the hypothesis that the distribution is one-parameter exponential and would accept the hypothesis that it is two-parameter exponential (see Figure 2).

Remark. One may also find an approximate estimate of A by looking at the graph. In our particular example, A is approximately equal to 150.

In example 8, we give a numerical procedure for testing whether $A > 0$, based on procedure 5.

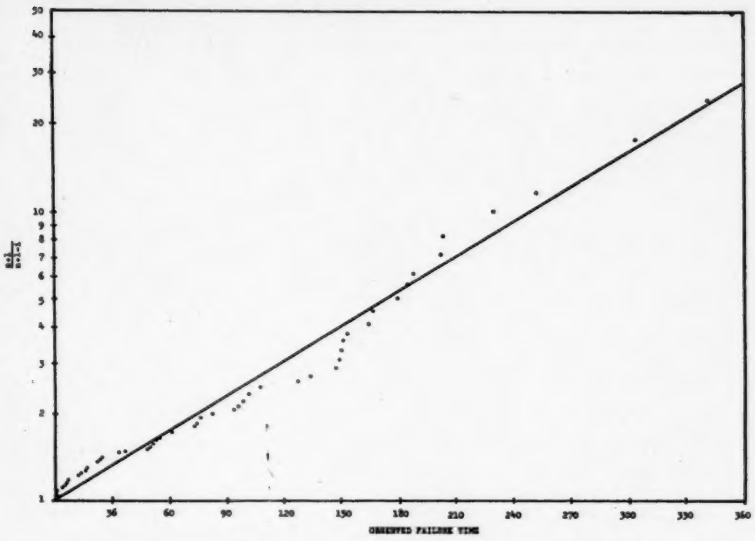


FIGURE 1—Graph of $y_i = \log (n + 1)/(n + 1 - i)$ against τ_i .
Data from Example 1 ($n = 49$).

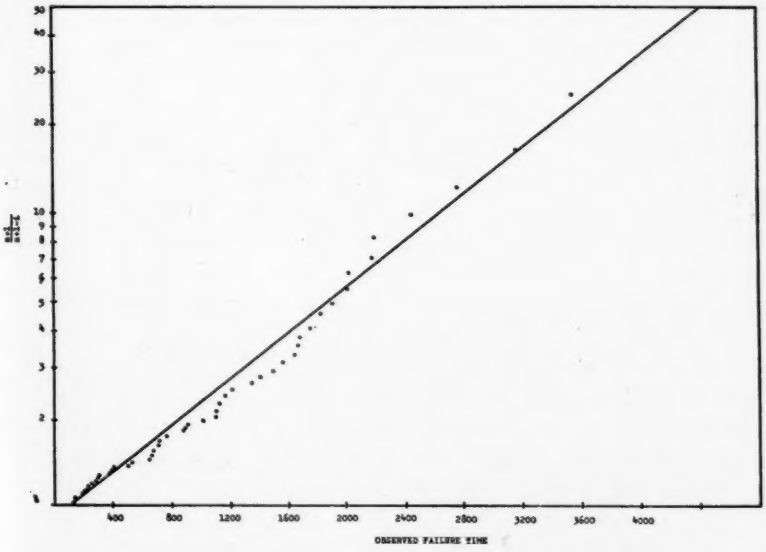


FIGURE 2—Graph of $y_i = \log (n + 1)/(n + 1 - i)$ against τ_i .
Data from Example 2 ($n = 49$).

(3) One observes a stochastic process until τ_{13}, the time when the 13th event occurs. The 13 observed times are $\tau_1 = 5.6$, $\tau_2 = 12.8$, $\tau_3 = 17.6$, $\tau_4 = 27.6$, $\tau_5 = 44.2$, $\tau_6 = 45.5$, $\tau_7 = 68.8$, $\tau_8 = 89.2$, $\tau_9 = 90.7$, $\tau_{10} = 100.2$, $\tau_{11} = 106.5$, $\tau_{12} = 115.5$, $\tau_{13} = 123.8$.

Are these observations consistent with the hypothesis that the underlying process is Poisson with constant rate parameter? Use procedure 3.

Solution: If the Poisson assumption holds, then the 12 times $\tau_1, \tau_2, \cdots, \tau_{12}$ should be tested to see whether they are uniformly distributed in $(0, \tau_{13})$. The theoretical expectation of $\sum_{i=1}^{12} \tau_i$ is $(12/2)\tau_{13} = 6(123.8) = 742.8$. The theoretical standard deviation is $(12/12)\ (123.8) = 123.8$. Under the hypothesis that the underlying process is Poisson with constant rate parameter, $\sum_{i=1}^{12} \tau_i$ is approximately normally distributed with mean 742.8 and standard deviation 123.8. Hence a 95% acceptance interval for the hypothesis being tested is $742.8 \pm (1.96)(123.8) = 742.8 \pm 242.6 = (500.2, 985.4)$. The observed sum is $\sum_{i=1}^{12} \tau_i = 724.2$. Since this falls within the acceptance interval, we accept the hypothesis that the underlying process is Poisson with constant rate parameter at the .05 level of significance.

(4) Twenty items are placed on test. The test is discontinued after 11 failures occur. Items are not replaced as they fail. The first 11 failure times are $\tau_1 = 7$, $\tau_2 = 12$, $\tau_3 = 15$, $\tau_4 = 24$, $\tau_5 = 25$, $\tau_6 = 48$, $\tau_7 = 53$, $\tau_8 = 56$, $\tau_9 = 72$, $\tau_{10} = 95$, $\tau_{11} = 100$. Are these data consistent with the hypothesis that the underlying distribution of life is exponential? Use procedure 3.

Solution: Let us compute $V(\tau_i)$ for $i = 1, 2, \cdots, 11$. Since the number of items originally on test is $n = 20$, $V(\tau_1) = 20(7) = 140$. $V(\tau_2) = V(\tau_1) + V(\tau_2 - \tau_1) = 140 + 19(5) = 140 + 95 = 235$. Proceeding in this way we get: $V(\tau_1) = 140$, $V(\tau_2) = 235$, $V(\tau_3) = 289$, $V(\tau_4) = 442$, $V(\tau_5) = 458$, $V(\tau_6) = 803$, $V(\tau_7) = 873$, $V(\tau_8) = 912$, $V(\tau_9) = 1104$, $V(\tau_{10}) = 1357$, $V(\tau_{11}) = 1407$. Under the assumption of exponentiality, the 10 total lives $V(\tau_i)$, $i = 1, 2, \cdots, 10$, should be uniformly distributed in $(0, V(\tau_{11}))$. To test this we calculate the theoretical mean and theoretical standard deviation for $\sum_{i=1}^{10} V(\tau_i)$. These are $(10/2)V(\tau_{11}) = (10/2)(1407) = 7035$ and $\sqrt{10/12}\ (1407) = (.913)(1407) = 1285$, respectively. Under the exponential assumption, $\sum_{i=1}^{10} V(\tau_i)$ is approximately normally distributed with mean 7035 and standard deviation 1285. Hence a 95% acceptance interval for the hypothesis that the underlying distribution is exponential is given by $7035 \pm 1.96(1285) = 7035 \pm 2518 = (4517,9553)$. But the observed sum is $\sum_{i=1}^{10} V(\tau_i) = 6613$. This is well within the acceptance interval and hence the hypothesis that the underlying distribution is exponential is accepted at the .05 significance level.

(5) In a *non-replacement* situation, a life test is terminated at a preassigned *total life* of 1000 hours. The *total lives* $T(\tau_1), \cdots, T(\tau_{45})$ are as follows:

30	173	266	346	468	611	662	763	915
63	178	281	392	486	616	692	773	917
99	207	290	433	562	631	693	776	918
102	244	310	436	563	651	693	807	964
172	255	343	455	584	660	700	844	978

where $T(\tau_1) = n\tau_1$ and $T(\tau_r) = \sum_{i=1}^{r-1} \tau_i + (n - r + 1)\tau_r$, $r = 2, \cdots, 45$.

Are these observations consistent with the hypothesis that the underlying process is Poisson with constant rate parameter? Use procedure 3.

Solution: Using procedure 3, it follows that if the Poisson assumption holds, then the total lives $T(\tau_1), \cdots, T(\tau_{45})$ are uniformly distributed in (0,1000). Since the sample is fairly large, the χ^2 goodness of fit test may be applied to see whether there are abnormal deviations from the uniform distribution. Dividing the interval (0,1000) into 5 equal intervals, we have*

	0–200	200–400	400–600	600–800	800–1000
Expected Number (e_i)	9	9	9	9	9
Observed Number (o_i)	7	10	8	13	7

$$\chi^2 = \sum_{i=1}^{5} \frac{(o_i - e_i)^2}{e_i} = \frac{26}{9} = 2.89.$$

Since we have divided the interval (0,1000) into 5 subintervals or classes, $\sum_{i=1}^{5} [(o_i - e_i)^2]/e_i$ is distributed as $\chi^2(4)$. But $Pr(\chi^2(4) > 2.89)$ lies between .5 and .7. Hence we accept the hypothesis that the underlying process is Poisson with constant rate parameter.

(6) In a life test we obtain the first 10 failure times. We suspect that the first 2 failures may have occurred unusually soon and should be rejected. The total life associated with the first two failures is $T(\tau_2) = 24$ and $T(\tau_{10})$, the total life associated with all ten failures, is 600. Test whether τ_1 and τ_2 occur abnormally soon. Carry out the test on the .05 significance level, using procedure 4.

Solution: Let us substitute in (11), using $\alpha = .05$. The lower 5% point of $F(4,16)$ is $F_\alpha^{**} = 1/5.84$, since $Pr(F(4,16) \leq 1/5.84) = .05$. Hence we will reject if

$$T(\tau_2) < \frac{2(600 - 24)}{(5.84)8},$$

i.e., if

$$T(\tau_2) < \frac{576}{23.36} = 24.7.$$

Since $T(\tau_2) = 24$, we reject the null hypothesis on the .05 significance level and instead have evidence that the first two failures occur abnormally soon.

* Note that we have used classes of *equal probability* rather than *equal length*. This removes the ambiguity of chi-square that was discussed in our description of procedure 2. For a discussion of this point, see Gumbel [20] and [21, p. 28]. Even when using classes of equal probability, one ambiguity remains. How many classes should be chosen? A reasonable rule is to divide the interval in such a way that the common value of the expected number $e_i \geq 5$.

(7) Twenty items are placed on life test. The life test is discontinued after 11 items have failed. Items that fail are not replaced. The first 11 failure times are $\tau_1 = 107$, $\tau_2 = 112$, $\tau_3 = 115$, $\tau_4 = 124$, $\tau_5 = 125$, $\tau_6 = 148$, $\tau_7 = 153$, $\tau_8 = 156$, $\tau_9 = 172$, $\tau_{10} = 195$, and $\tau_{11} = 200$. The first failure time appears to be abnormally long and one wishes to test whether the underlying distribution is $(1/\theta)e^{-(x-A)/\theta}$, with $A = 0$ against the alternative that $A > 0$. Carry out the test at the .01 significance level, using procedure 5.

Solution: It can be computed readily that $T(\tau_1) = 2140$ and $T(\tau_{11}) = 3407$. Hence $T(\tau_{11} - \tau_1) = 3407 - 2140 = 1267$. Substituting in (13) with $\alpha = .01$, we see that we will reject the hypothesis that $A = 0$, if $T(\tau_1) > (5.85)(1267)/10 = 741$ [5.85 is the upper 1% point of the $F(2,20)$ distribution].

But $T(\tau_1)$ was observed to equal 2140. Hence we reject the hypothesis that $A = 0$ at the .01 significance level. We have strong grounds for believing that the minimum life $A > 0$.

(8) Test whether $A > 0$, i.e., whether the first failure time is abnormally long in example 2. Carry out the test on the .01 significance level using procedure 5.

Solution: In this problem $T(\tau_1) = 7938$ and $T(\tau_{49}) = 58432$. Hence $T(\tau_{49} - \tau_1) = 58432 - 7938 = 50{,}494$. We will reject the hypothesis that $A = 0$ on the .01 significance level if

$$T(\tau_1) > G^*_{.01} \frac{T(\tau_{49} - \tau_1)}{48},$$

where $G^*_{.01} = 4.83$ is the upper 1% point of the $F(2, 96)$ distribution. Since the right-hand side is 5050 and $T(\tau_1) = 7938$, we reject the hypothesis that $A = 0$ at the 1% significance level. There is good reason to suspect that it took abnormally long for the first failure to occur.

(9) A life test is carried out and one is given the first twenty failure times $\tau_1 \leq \tau_2 \leq \cdots \leq \tau_{20}$. The total life $T(\tau_{10})$ associated with the time interval $[0, \tau_{10}]$ is 10,000 hours and the total life $T(\tau_{20} - \tau_{10})$ associated with the time interval $[\tau_{10}, \tau_{20}]$ is 25,000 hours. Let us define θ_1 as the mean life in $[0, \tau_{10}]$ and θ_2 as the mean life in $[\tau_{10}, \tau_{20}]$. (a) Test the hypothesis $\theta_1 = \theta_2$ against the two-sided alternative that $\theta_1 \neq \theta_2$, at the .05 significance level; (b) test the hypothesis that $\theta_1 = \theta_2$ against the one-sided alternative that $\theta_2 > \theta_1$, at the .05 significance level. Use procedure 6.

Solution: Using procedure 6, let us note that $2T(\tau_{10})/\theta_1$ and $2T(\tau_{20} - \tau_{10})/\theta_2$ are mutually independent and are each distributed as $\chi^2(20)$. If, in particular, the null hypothesis is satisfied, $\theta_1 = \theta_2 = \theta$ and so $T(\tau_{20} - \tau_{10})/T(\tau_{10})$ is distributed as $F(20, 20)$.

But $Pr(1/2.46 < F(20, 20) < 2.46) = .95$. Hence if we wish to have a two-sided test with type 1 error = .05, we should accept the hypothesis that $\theta_1 = \theta_2$ if

$$\frac{1}{2.46} < \frac{T(\tau_{20} - \tau_{10})}{T(\tau_{10})} < 2.46$$

and reject otherwise. The observed value of

$$\frac{T(\tau_{20} - \tau_{10})}{T(\tau_{10})} = \frac{25{,}000}{10{,}000} = 2.50.$$

Hence we reject the hypothesis that $\theta_1 = \theta_2$ on the .05 level.

In the one-sided case, we note that $Pr[F(20, 20) > 2.12] = .05$. Hence we will reject the hypothesis that $\theta_1 = \theta_2$ (and accept $\theta_2 > \theta_1$) if

$$\frac{T(\tau_{20} - \tau_{10})}{T(\tau_{10})} > 2.12.$$

Thus we reject the hypothesis that $\theta_1 = \theta_2$ on the .05 level, since we observed $T(\tau_{20} - \tau_{10})/(\tau_{10}) = 2.5$.

Remark: In actual practice there may well be *a priori* grounds for using either the one- or two-sided test. Suppose we know, e.g., that the quality (as measured by length of life) of the last 10 items tested is at least as good as the quality of the first 10 items. If that is the case, a one-sided test should be used.

(10) One observes a life test until τ_{51}, the time when the 51st failure occurs. The observed total lives are:

150	1220	1961	2476	3270	3575	3790
314	1273	2165	2789	3297	3577	3799
493	1329	2180	2838	3348	3577	3810
517	1401	2275	2853	3357	3579	
858	1449	2282	3006	3402	3621	
865	1549	2434	3191	3461	3656	
889	1716	2439	3228	3501	3694	
1033	1728	2441	3238	3509	3760	

It is suspected that as time goes on, there is a deterioration in quality so that the mean life θ decreases as a function of time. Hence we are interested in testing the hypothesis that the observations are coming from an exponential distribution with constant parameter θ against the alternative that the parameter θ decreases during the life test. Use procedure 6 at the .05 significance level.

Solution: In using procedure 6 we shall be testing a somewhat simpler hypothesis. Let us define θ_1 as the mean life in $[0, \tau_{25}]$ and θ_2 as the mean life in $[\tau_{25}, \tau_{51}]$. Then we should like to test the null hypothesis: $\theta_1 = \theta_2$ against the alternative: $\theta_2 < \theta_1$. Using procedure 6, $2T(\tau_{25})/\theta_1$ and $2T(\tau_{51} - \tau_{25})/\theta_2$ are mutually independent and distributed as $\chi^2(50)$ and $\chi^2(52)$, respectively. If the null hypothesis is satisfied, $\theta_1 = \theta_2 = \theta$ and the statistic $[25T(\tau_{51} - \tau_{25})]/26T(\tau_{25})$ is distributed as $F(52, 50)$. The null hypothesis will be rejected for sufficiently small values of F. Using the given data, we have

$$\frac{25T(\tau_{51} - \tau_{25})}{26T(\tau_{25})} = \frac{25(3810 - 2476)}{26(2476)} = .518.$$

From the F table we have that $Pr[F(52, 50) < 1/1.60 = .625] = .05$. But $.518 < .625$, hence we reject the hypothesis that $\theta_1 = \theta_2$ at the .05 significance

level and assert that θ_2 , the mean life in the second half of the experiment, is less than θ_1 , the mean life in the first half of the experiment.

Remark: A graphical way of seeing whether θ changes in time is to plot $T(\tau_{i+1} - \tau_i)$, the elapsed life between successive failures, against i. An analytical procedure is to compute a rank correlation coefficient for the $\{T(\tau_{i+1} - \tau_i)\}$.

(11) We are given the results of 5 life tests in each of which 10 failures were observed. The total lives in each of the life tests are 12.74, 10.02, 7.25, 9.18, and 12.99 (thousand hours), respectively. Use procedures 7 or 9 and the formulae in Appendix 2 to test whether θ is constant over all the tests.

Solution: Let us carry out the solution using Appendix 2. In this problem r, the number of failures in each group, is 10 and k, the number of groups, is 5. It is readily computed that

$$\log L_1 = \tfrac{1}{5}[\log 12.74 + \log 10.02 + \log 7.25 + \log 9.18 + \log 12.99] - \log\left(\frac{52.18}{5}\right)$$

$$= 2.3223 - 2.3453 = -.0230$$

$$\therefore L_1 = .9772.$$

Reference to the L_1 tables for $n = 2r + 1 = 21$ and $k = 5$, indicates that the hypothesis that θ is constant should be rejected if $L_1 < .905$, if we wish to maintain a Type I error = .05. The value actually obtained is .9772 and so the hypothesis is accepted on the .05 level. Another way of carrying out procedure 7 involves computing

$$\frac{-2rk \log L_1}{c}, \quad \text{where} \quad c = 1 + \frac{k+1}{6rk} \quad \text{and where} \quad \frac{-2rk \log L_1}{c}$$

is distributed as $\chi^2(4)$ (see formula (17) in procedure 7 and the discussion in Appendix 3). Since $r = 10$, $k = 5$, we get $c = 1 + 1/50 = 51/50$ and

$$\frac{-2rk \log L_1}{c} = \left(\frac{50}{51}\right)(2.30) = 2.26.$$

If we wish to have the Type I error = .05, the hypothesis of homogeneity will be rejected if

$$\frac{-2rk \log L_1}{c} > 9.49,$$

the upper 5% point of the $\chi^2(4)$ distribution. But the observed value of

$$\frac{-2rk \log L_1}{c} = 2.26.$$

Hence the hypothesis is accepted. An alternative test for homogeneity is provided by procedure 9. This involves computing

$$U = \max T_i / \min T_i .$$

In this case $U = 12.99/7.25 = 1.79$.

We enter Hartley's table 3 (p. 310) for $k = 5$, $\nu = 20$. The upper 5% point for U is 3.49 (i.e., if we reject the hypothesis of homogeneity for $U > 3.49$ and accept if $U \leq 3.49$, our Type I error = .05). But we observed $U = 1.79$. Hence the hypothesis of homogeneity is accepted.

(12) Five items are observed to fail on a life test. The total lives between successive failures (first total life is measured from $t = 0$, when the life test started) are 627, 898, 830, 50, and 1028. Use procedures 8 or 9 and Appendix 2 to test whether θ is constant throughout. Use the 5% level of significance in making computations.

Solution: Using formulae in Appendix 2, one gets

$$\log L_1 = 1/5[\log 627 + \log 898 + \log 830 + \log 50 + \log 1028] - \log\left(\frac{3433}{5}\right)$$

$$= 6.1620 - 6.5317 = -.3697.$$

$$\therefore L_1 = .691.$$

Reference to the L_1 tables for $n = 3$ and $k = 5$ indicates that the hypothesis that θ is a constant should be accepted if $L_1 < .328$, if we wish to maintain a Type I error = .05. The value actually obtained is .691 and so the hypothesis is accepted at the .05 level of significance. Another way of carrying out procedure 8 involves computing

$$\frac{-2k \log L_1}{c},$$

where

$$c = 1 + \frac{k+1}{6k}.$$

Since $k = 5$, we get $c = 6/5$ and

$$\frac{-2k \log L_1}{c} = \frac{5}{6}(3.697) = 3.081.$$

If we wish to have the Type I error = .05, the hypothesis of homogeneity will be rejected if

$$\frac{-2k \log L_1}{c} > 9.49$$

(since $(-2k \log L_1)/c$ is distributed as $\chi^2(4)$ and the upper 5% point of $\chi^2(4)$ is 9.49). But the observed value of

$$\frac{-2k \log L_1}{c} = 3.08.$$

Hence the hypothesis that θ is constant throughout is accepted.

Hartley's test based on the maximum F distribution (procedure 9) gives

$$U = \frac{\max T_i}{\min T_i} = \frac{1028}{50} = 20.56.$$

Since $Pr(U \geq 202) = .05$, we are led to accept the hypothesis that θ is constant throughout.

(13) A life test is carried out and one is given the total lives associated with the first twenty failure times. The total lives $T(\tau_1), \cdots, T(\tau_{20})$ are

16	34	77	131
18	55	95	144
23	57	120	157
26	67	121	195
27	74	129	198

Test the hypothesis that θ is constant in $(0, \tau_{20})$ against the alternative that it varies in $(0, \tau_{20})$. Use procedure 7 (Bartlett's test), by splitting the 20 observations into 4 groups of 5 observations each.

Solution: If the underlying process is Poisson, then

$$\frac{2rk\left\{\log\left(\frac{T(\tau_k)}{k}\right) - \frac{1}{k}\left[\log T(\tau_{1r}) + \log T(\tau_{2r} - \tau_{1r}) + \cdots + \log T(\tau_{kr} - \tau_{(k-1)r})\right]\right\}}{1 + \frac{k+1}{6rk}}$$

is approximately distributed as χ^2_{k-1}. In our example the number of groups is $k = 4$ and the number of items in a group is $r = 5$. Therefore, we have

$$B = \frac{2(5)(4)\left\{\log \frac{198}{4} - \frac{1}{4}\left[\log 27 + \log (74 - 27)\right.\right.}{1 + \frac{4+1}{6(5)(4)}}$$

$$\cdot \frac{\left.\left. + \log (129 - 74) + \log (198 - 129)\right]\right\}}{1 + \frac{4+1}{6(5)(4)}} = 2.20$$

From the χ^2 table we have that $P\{\chi^2_3 < 7.82\} = .95$. Hence we accept the hypothesis that θ is constant at the .05 significance level.

(14) Five items are observed to fail on a life test. The total lives between successive failures (first total life is measured from $t = 0$, when the life test started) are 68, 1516, 49, 22, and 358. Use Fisher's test (procedure 10) to see whether $T(\tau_2 - \tau_1) = 1516$ is abnormally long. Do this at the 5% and 1% significance levels.

Solution: According to (20), we compute

$$Z = \frac{\max \{T(\tau_1), T(\tau_2 - \tau_1), \cdots, T(\tau_5 - \tau_4)\}}{T(\tau_5)}.$$

In this problem

$$Z = 1516/2013 = .753.$$

According to a table given by Fisher:

$$Pr(Z \geq .684) = .05 \text{ when the number of failures is } n = 5.$$

But we observed $Z = .753$. Hence there is evidence at least on the .05 level that $T(\tau_2 - \tau_1) = 1516$ is abnormally long. It is interesting to note that $Pr(Z \geq .789) = .01$, hence we would not be able to assert that $T(\tau_2 - \tau_1)$ is abnormally long on the .01 level. The chances are better than one in a hundred of observing a $T(\tau_2 - \tau_1)$ longer than the one actually obtained.

(15) One observes a life test (stochastic process) until τ_{21}, the time when the 21st failure (event) occurs. The observed total lives are:

$$T(\tau_1) = 13, \quad T(\tau_2) = 23, \quad T(\tau_3) = 44, \quad T(\tau_4) = 75, \quad T(\tau_5) = 81,$$
$$T(\tau_6) = 83, \quad T(\tau_7) = 122, \quad T(\tau_8) = 164, \quad T(\tau_9) = 166, \quad T(\tau_{10}) = 182,$$
$$T(\tau_{11}) = 218, \quad T(\tau_{12}) = 247, \quad T(\tau_{13}) = 289, \quad T(\tau_{14}) = 334,$$
$$T(\tau_{15}) = 361, \quad T(\tau_{16}) = 373, \quad T(\tau_{17}) = 377, \quad T(\tau_{18}) = 426,$$
$$T(\tau_{19}) = 447, \quad T(\tau_{20}) = 451, \quad T(\tau_{21}) = 526.$$

Are these observations consistent with the hypothesis that the underlying process is Poisson with a constant rate parameter? Carry out a two-sided Kolmogorov-Smirnov test (procedure 11) at the .05 significance level.

Solution: Graphical Method (see Figure 3): If the underlying process is Poisson with constant rate parameter, then the $T(\tau_i)$, $i = 1, 2, \cdots, 20$ are uniformly distributed over the interval $[0, T(\tau_{21})]$, i.e., over $[0, 526]$. The theoretical c.d.f. is given by the equation

$$\begin{aligned} F(T) &= 0, && T \leq 0 \\ &= T/526, && 0 \leq T \leq 526 \\ &= 1, && T \geq 526. \end{aligned}$$

The theoretical c.d.f. is represented graphically in Figure 3 by the line L_1. The sample c.d.f. associated with the data is given by $F_n(T)$, which is defined as follows:

$$\begin{aligned} F_n(T) &= 0, \quad \text{for} \quad T \leq 0 \\ &= \frac{j-1}{20} \quad \text{for} \quad T(\tau_{j-1}) \leq T(\tau) < T(\tau_j), \qquad j = 1, 2, \cdots, 20 \\ &= 1, \quad \text{for} \quad T(\tau) \geq T(\tau_{20}) = 451. \end{aligned}$$

Put simply, $F_n(T)$ is a step function, with jumps of $1/20$ occurring at the total lives $T(\tau_j)$, $j = 1, 2, \cdots, 20$ associated with the failure times τ_j.

To test the hypothesis that the underlying process is Poisson with constant rate parameter, we wish to see whether $F_n(T)$ differs "too much" from $F(T)$. Using the table of the Kolmogorov statistic for the 95% point when the sample size is $n = 20$ (see [10]), we can state that if the null hypothesis is true, then there is probability .95 that the maximum deviation between $F_n(T)$ and $F(T)$ is in absolute value $\leq .2939$. Graphically this means the following:

Draw two lines, L_2 and L_3, parallel to L_1 such that the vertical distance

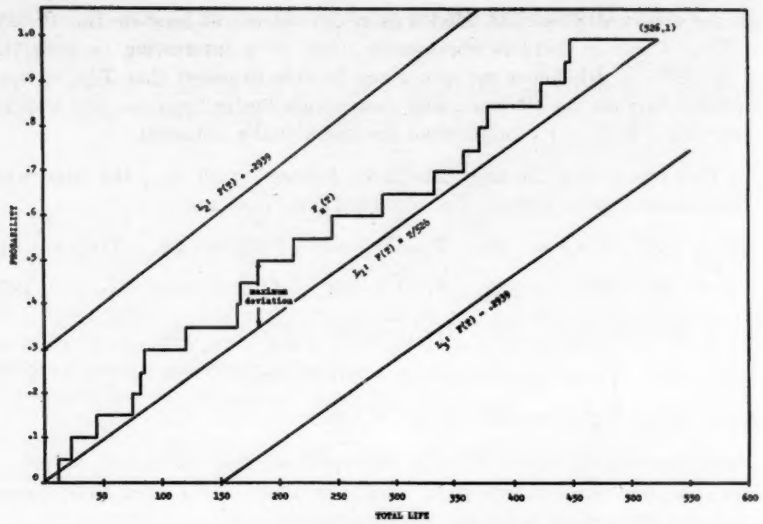


FIGURE 3—The Kolmogorov-Smirnov test carried out graphically for the data in Example 15.

between them and L_1 is $\pm.2939$. Rejection of the null hypothesis at the .05 significance level will occur if the step function $F_n(T)$ crosses either L_2 or L_3. If $F_n(T)$ does not cross either of the boundary lines, we will accept. In this problem $F_n(T)$ does not cross either L_2 or L_3 and hence we accept the hypothesis that the underlying process is Poisson with a constant rate parameter.

ANALYTICAL METHOD FOR PERFORMING THE KOLMOGOROV-SMIRNOV TEST

An equivalent analytical procedure, which is sometimes more convenient, is to compute

$$D_n = \text{l.u.b.}\ |F(T) - F_n(T)|$$

directly. In our particular problem, this becomes

$$D_{20} = \max_{1 \le i \le 21} \left[\max \left\{ \left| \frac{T(\tau_j)}{526} - \left(\frac{j-1}{20} \right) \right|, \left| \frac{T(\tau_{j-1})}{526} - \left(\frac{j-1}{20} \right) \right| \right\} \right].$$

It is readily calculated that $D_{20} = .154$.* But from the table of the Kolmogorov statistic [10], we have that $Pr(D_{20} \le .2939) = .95$. Therefore, we accept the hypothesis that the underlying process is Poisson with constant parameter.

Remark: Still another possible solution (see Remark 2 following procedure 8) is based on the fact that

$$-2 \sum_{i=1}^{r-1} \log \left[\frac{T(\tau_i)}{T(\tau_r)} \right] = -2 \left[\sum_{i=1}^{r} \log T(\tau_i) - r \log T(\tau_r) \right]$$

* To compute D_{20}, we readily verify that the maximum deviation is attained at the 10th observation. At time $T(\tau_{10}) = 182$, $F_n(T) - F(T) = .5 - 182/526 = .5 - .346 = .154$.

is distributed as χ^2_{2r-2}, when the assumption is made that the underlying process is Poisson. In the present problem it can be computed that

$$-2\left[\sum_{i=1}^{21} \log T(\tau_i) - 21 \log 526\right] = 48.18.$$

From the table of χ^2 we have that

$$Pr(24.4 < \chi^2_{40} < 59.3) = .95$$

and hence if we use the .05 significance level, we accept the hypothesis that the rate parameter is constant, when it is tested against the two-sided alternative that it is not constant. If a one-sided test had been run, we would still have accepted the hypothesis that the Poisson rate parameter is constant, when it is tested against the alternative that the Poisson rate is decreasing in time. This is so because the upper 5% point of χ^2_{40} is 55.76 and we observed 48.18.

Although the null hypothesis is accepted by both tests at the 5% significance level, the procedure based on

$$-2\sum_{i=1}^{r-1} \log\left[\frac{T(\tau_i)}{T(\tau_r)}\right]$$

is more sensitive. That this should be the case is made clear from considerations in [4] and [5].

(16) A non-replacement life test is terminated at a preassigned total life of 4100 hours. The total lives $T(\tau_1), \cdots, T(\tau_{80})$ are:

32	151	423	1602	2611	3581	3883	4001
35	169	587	1674	2701	3733	3884	4024
45	194	766	1722	3084	3738	3907	4040
52	195	790	1822	3089	3740	3914	4043
54	203	1131	1989	3192	3776	3929	4069
68	205	1138	2009	3326	3807	3933	4070
109	218	1162	2234	3351	3812	3946	4070
114	231	1306	2438	3428	3814	3957	4071
133	269	1493	2458	3496	3829	3977	4088
148	272	1546	2548	3574	3848	3981	4100

Test the hypothesis that the underlying process is Poisson with parameter θ against the alternative that the parameter θ has changed. Use procedures 3, 7, or 9, all at the .05 significance level.

Solution: In procedure 3 we use the fact that if the Poisson assumption holds, then the total lives $\{T(\tau_1), \cdots, T(\tau_{80})\}$ are uniformly distributed in [0,4100]. Because the sample is fairly large, it is convenient to apply the χ^2 goodness of fit test, where we are testing for deviations from the null hypothesis that observations are uniformly distributed over the interval [0,4100]. Dividing the interval [0,4100] into 10 equal intervals we have

	0–410	410–820	820–1230	1230–1640	1640–2050	2050–2460	2460–2870	2870–3280	3280–3690	3690–4100
Expected number (e_i)	8	8	8	8	8	8	8	8	8	8
Observed number (o_i)	20	4	3	4	5	3	3	3	6	29

The calculated value of χ^2 is

$$\chi^2 = \sum_{i=1}^{10} \frac{(o_i - e_i)^2}{e_i} = \frac{730}{8} = 91.25.$$

From the χ^2 table we have that $Pr[\chi_9^2 > 16.9] = .05$. Hence we reject the null hypothesis at the .05 significance level. In procedure 7, we have a choice as to how we should subdivide our data into groups. For convenience we choose $k = 4$ groups containing $r = 20$ observations each. That is, we are interested in the total lives associated with the time intervals $[0, \tau_{20}]$, $[\tau_{20}, \tau_{40}]$, $[\tau_{40}, \tau_{60}]$, $[\tau_{60}, \tau_{80}]$. The associated total lives are: $T(\tau_{20}) = 272$, $T(\tau_{40} - \tau_{20}) = 2548 - 272 = 2276$, $T(\tau_{60} - \tau_{40}) = 3848 - 2548 = 1300$, and $T(\tau_{80} - \tau_{60}) = 4100 - 3848 = 252$. Substituting into (16), using $k = 4$, $r = 20$, we get

$$B = \frac{2(20)(4)\left\{\log\left(\frac{4100}{4}\right) - \frac{1}{4}\left[\log 272 + \log 2276 + \log 1300 + \log 252\right]\right\}}{1 + \frac{4+1}{6(20)(4)}} = 67.07.$$

From the χ^2 table we have that $Pr[\chi_3^2 > 7.81] = .05$. Hence we reject the hypothesis that θ is constant. In procedure 9, we apply Hartley's U test, i.e., we compute

$$U = \frac{\max\{T(\tau_{20}), T(\tau_{40} - \tau_{20}), T(\tau_{60} - \tau_{40}), T(\tau_{80} - \tau_{60})\}}{\min\{T(\tau_{20}), T(\tau_{40} - \tau_{20}), T(\tau_{60} - \tau_{40}), T(\tau_{80} - \tau_{60})\}}$$

$$= \frac{2276}{252} = 9.03.$$

Entering Hartley's table for $k = 4$, $\nu = 40$, we see that the upper 5% point of U is about 2.35. (The table does not give the upper 5% point for $k = 4$, $\nu = 40$. It does give 2.59 as the upper 5% point for $k = 4$, $\nu = 60$.) Hence the hypothesis that θ is constant is rejected.

Remark: The deviations from the null hypothesis are so extreme in this problem that one scarcely needs a statistical test.

(17) A life test is terminated at a preassigned total life of 1000 hours. The total lives $T(\tau_1), \cdots, T(\tau_{40})$ are:

32	68	151	205	315	412	625	832
35	109	169	218	326	455	700	891
45	114	194	231	327	480	704	948
52	133	195	269	327	555	757	957
54	148	203	272	399	574	810	985

Test the hypothesis that the process is Poisson with constant rate parameter against the alternative that the value of θ increases during the test. Try procedures 3, 7, and 11. Test at the .05 significance level.

Solution: Procedure 3 is based on the result that if the Poisson assumption holds, then the total life is uniformly distributed in (0,1000). Since the sample is fairly large, the χ^2 goodness of fit criterion may be used to test uniformity. Dividing the interval (0,1000) into 5 equal intervals we have:

	0–200	200–400	400–600	600–800	800–1000
Expected number (e_i)	8	8	8	8	8
Observed number (o_i)	14	11	5	4	6

$$\chi^2 = \sum_{i=1}^{5} \frac{(o_i - e_i)^2}{e_i}$$

is distributed as χ_4^2. In this problem

$$\sum_{i=1}^{5} \frac{(o_i - e_i)^2}{e_i} = \frac{74}{8} = 9.25.$$

Since $Pr\{\chi_4^2 > 9.49\} = .05$, we just barely accept the hypothesis that θ is constant over the total life interval [0,1000].

Remark: When we applied this test, we did not use the fact that, for the alternative stated, we were testing against an increase in the parameter. This suggests that an appropriate test might be to compute a rank correlation coefficient for the $T(\tau_i - \tau_{i-1})$, $i = 1, 2, \cdots, 40$.

A test which is more sensitive in this situation is Bartlett's test (procedure 7). From previous considerations we have that if the underlying process is Poisson, then

$$B = \frac{2rk\left\{\log \frac{T(\tau_{kr})}{k} - \frac{1}{k}\left[\log T(\tau_{1r}) + \sum_{i=2}^{k} \log T(\tau_{ir} - T_{(i-1)r})\right]\right\}}{1 + \frac{k+1}{6rk}}$$

is distributed as χ_{k-1}^2.

In our example if we let the number of items in a group be $r = 8$, and the number of groups be $k = 5$, we have that

$$B = \frac{2(8)(5)\left\{\log\left(\frac{985}{5}\right) - \frac{1}{5}\left[\log 114 + \log (205 - 114)\right.\right.}{1 + \frac{6}{6(5)(8)}}$$

$$\frac{\left.\left. + \log (327 - 205) + \log (700 - 327) + \log (985 - 700)\right]\right\}}{1 + \frac{6}{6(5)(8)}}$$

$$= 12.346.$$

Since we are testing against the alternative that θ increases, we have a one-sided test. We reject the hypothesis that θ is constant because from the χ^2 table we have that

$$P\{\chi_4^2 > 9.49\} = .05.$$

Finally, we apply the Kolmogorov-Smirnov test (procedure 11). Since the life test is terminated at a preassigned total life of 1000 hours, the 40 total lives $\{T(\tau_i)\}$, $i = 1, 2, \cdots, 40$ which occurred prior to total life = 1000 hours are uniformly distributed over the interval [0,1000], if the underlying process is Poisson with constant mean life θ. The theoretical c.d.f. is defined as:

$$\begin{aligned} F(T) &= 0, && T \leq 0 \\ &= T/1000, && 0 \leq T \leq 1000 \\ &= 1, && T \geq 1000. \end{aligned}$$

The theoretical c.d.f. is represented graphically in Figure 4 by the line L_1. The sample c.d.f. associated with the data is given by $F_n(T)$, which is defined as follows:

$$\begin{aligned} F_n(T) &= 0, && \text{for } T \leq 0 \\ &= \frac{j-1}{40}, && \text{for } T(\tau_{j-1}) \leq T(\tau) < T(\tau_j), \quad j = 1, 2, \cdots, 40 \\ &= 1, && \text{for } T(\tau) \geq T(\tau_{40}) = 985. \end{aligned}$$

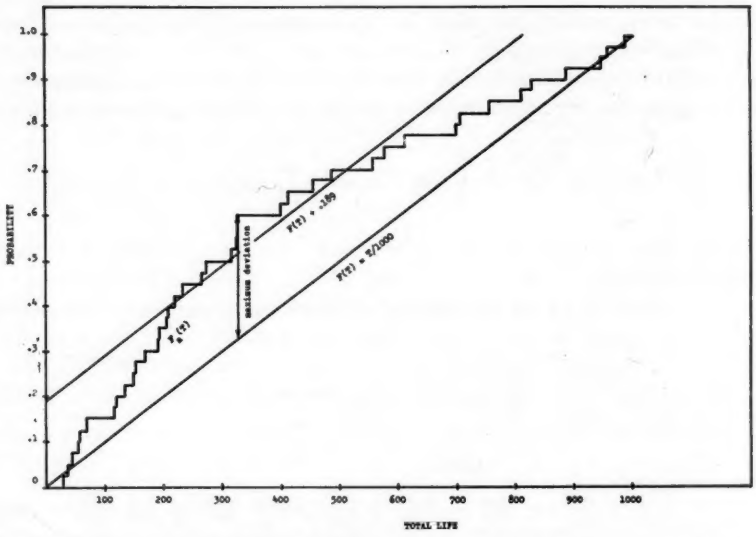


FIGURE 4—The one-sided Kolmogorov-Smirnov test carried out graphically for the data in Example 17.

Put simply, $F_n(T)$ is a step function, with jumps of 1/40 occurring at the total lives $T(\tau_j)$, $j = 1, 2, \cdots, 40$.

Since we are testing the null hypothesis that θ is constant against the one-sided alternative that θ increases in time, we will use a one-sided Kolmogorov-Smirnov type test. From table 1 in [9] we know that

$$\Pr\{\max_{0 \le T \le 1000} (F_n(t) - F(t)) > .189\} = .05,$$

when the sample size $n = 40$. Graphically this means the following: draw a line L_2, which is parallel to L_1 and is a vertical distance .189 above L_1. Rejection of the null hypothesis at the .05 level will occur if the step function $F_n(T)$ crosses L_2 (for, if this happens, $\max_{0 \le T \le 1000} (F_n(T) - F(T))$ does exceed .189). In this example $F_n(T)$ crosses L_2 and hence we reject the hypothesis that the mean life θ is constant throughout the interval [0,1000]. Actually,

$$\max_{1 \le i \le 40} \{F_n(T) - F(T)\} = \max_{1 \le i \le 40} \left\{\frac{j}{40} - \frac{T(\tau_j)}{1000}\right\}$$

$$= \left\{\frac{24}{40} - \frac{327}{1000}\right\} = .600 - .327 = .273$$

which exceeds .189. As a matter of fact, the 1% point of $\max_{0 \le T \le 1000} (F_{40}(T) - F(T))$ is .235 and so the null hypothesis would be rejected even at the .01 level.

For references and descriptions of the theory underlying the above examples, the reader is referred to Part I of this paper which appears in the previous issue of this journal: *Technometrics*, Vol. 2, No. 1, February 1960, p. 83–101.

Partial Duplication of Response Surface Designs

O. DYKSTRA, JR.

General Foods Research Center

Continuing the author's earlier work [6] a method is described which requires that certain experimental runs of a central composite, second-order, response surface design be repeated, thereby providing a more general estimate of the experimental error, at the same time providing more reliable estimates of the effects.

The partial duplication of the factorial portion as well as the partial duplication of the star portion has been considered and described. The response surface designs with the star portion duplicated seem to have more potential than the designs with their factorial portions duplicated or partially duplicated.

1. The Argument for Partial Duplication in Response Surface Designs

Many, perhaps most, experiments are based on prior knowledge, and what is desired is knowledge about the changes in response which occur when the present process is changed from some "normal" set of levels. In these cases the experimental design is centered according to the experimenter's best guess. Using the presently available second order designs, this "center position" is the condition at which the n_0 replicates are usually obtained. It is frequently feared that the variability may increase away from the center point, the resulting estimate of error may be too small for a correct assessment of the coefficients of the second order polynomial. Thus, since we often have no assurance that the experimental error is constant over the entire experimental area, it would be sound experimental strategy to obtain duplicates over the experimental area. Furthermore, such a study might reveal the existence of heterogeneous error.

Partial duplication requires additional runs, so that we also obtain more precision in the estimates of the coefficients, more degrees of freedom for an estimate of the experimental error, and a more powerful test of the adequacy of the second order model.

2. Response Surface Designs (Second-Order)

Before entering into a discussion of partial duplication of second-order response surface designs, it is necessary to review certain aspects of the work of Box, Hunter, et al [1, 2, 3, 4, and 5].

The number of variables is denoted by k. For the case of central composite

designs (to which attention in this paper is restricted), the total number of runs is $N = n_c + n_a + n_0$, where

$n_c = 2^k$ is the number of points in a square, cube, hypercube, or fraction of the hypercube. These points have coodinates $(\pm 1, \pm 1, \cdots, \pm 1)$. For $k \geq 5$ we may take $n_c = 2^{k-1}$, and for $k \geq 8$ we may take $n_c = 2^{k-2}$.

$n_a = 2k$ is the number of points in a "star" design. These points have coordinates $(\pm\alpha, 0, \cdots, 0)$, $(0, \pm\alpha, 0, \cdots, 0)$, $\cdots$, $(0, \cdots, 0, \pm\alpha)$.

n_0 is the number of points at the center of the design, i.e., with coordinates $(0, \cdots, 0)$.

With α chosen equal to $n_c^{\frac{1}{4}}$ the design becomes "rotatable", that is, the variance of a prediction, $V(\hat{y})$, is constant for all points the same distance, ρ, from the center of the design.

For the central composite designs each of the k variables are at five levels: $0, \pm 1$, and $\pm\alpha$. These settings are, of course, a convenient coding of the original variables. The n_0 points at the center of the design provide an estimate of the experimental error s^2 with $n_0 - 1$ degrees of freedom, which can be used for statistical tests of significance. These center points serve another important function, since they also greatly affect the value of $V(\hat{y})$ when ρ is not too large. Box and Hunter [2] give values of n_0 which force $V(\hat{y})$ for $\rho = 0$ to equal $V(\hat{y})$ when $\rho = 1$. Such designs are called "uniform information" designs.

DeBaun [5] and Box and Hunter [2] give methods for blocking second order designs, if blocking is necessary. In order to obtain orthogonal blocking, i.e., designs for which the block effects are independent of the estimates of the coefficients, it is necessary to allot the n_0 center points to the blocks so that the total number of points in the blocks is proportional to the respective values of $\sum x^2$ for each block.

The true response η is some function of the k variables $x_1, \cdots, x_k$; in general $\eta = \theta(x_1, \cdots, x_k)$. It is assumed here that this function can be adequately represented by a polynomial function which is at most quadratic in the area of interest, i.e.,

$$\eta = \beta_0 x_0 + \beta_1 x_1 + \cdots + \beta_k x_k + \beta_{11} x_1^2 + \cdots + \beta_{kk} x_k^2 + \beta_{12} x_1 x_2 + \cdots + \beta_{k-1,k} x_{k-1} x_k . \tag{1}$$

The value of x_0 is identically equal to 1. For a given set of conditions the response, y_u, has mean η_u. In general $y = \eta + \epsilon$, where ϵ are $NID(0, \sigma^2)$.

For the designs under consideration in this paper the matrix of sums of products and cross-products, $\mathbf{X'X}$ may be written

$$
(2)\quad \mathbf{X'X} = \begin{bmatrix}
A & 0 & 0 & \cdots & 0 & B & B & \cdots & B & 0 & 0 & \cdots & 0 \\
 & B & 0 & \cdots & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & \cdots & 0 \\
 & & B & \cdots & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & \cdots & 0 \\
 & & & \ddots & \vdots & \vdots & \vdots & & \vdots & \vdots & \vdots & & \vdots \\
 & & & & B & 0 & 0 & \cdots & 0 & 0 & 0 & \cdots & 0 \\
 & & & & & C & D & \cdots & D & 0 & 0 & \cdots & 0 \\
 & & & & & & C & \cdots & D & 0 & 0 & \cdots & 0 \\
 & & & & & & & \ddots & \vdots & \vdots & \vdots & & \vdots \\
 & & & & & & & & C & 0 & 0 & \cdots & 0 \\
 & & & & \text{(symmetric)} & & & & & D & 0 & \cdots & 0 \\
 & & & & & & & & & & D & \cdots & 0 \\
 & & & & & & & & & & & \ddots & \vdots \\
 & & & & & & & & & & & & D
\end{bmatrix}
\begin{matrix}
(x_0) \\ (x_1) \\ (x_2) \\ \vdots \\ (x_k) \\ (x_1^2) \\ (x_2^2) \\ \vdots \\ (x_k^2) \\ (x_1x_2) \\ (x_1x_3) \\ \vdots \\ (x_{k-1}x_k)
\end{matrix}
$$

Assuming that the variables take on values 0, ± 1, and $\pm\alpha$, then

$$
\begin{aligned}
A &= \sum x_0^2 = N = n_c + n_a + n_0 = D + 2k + n_0 \\
B &= \sum_u x_{iu}^2 = D + 2\alpha^2 \quad \text{for all} \quad i \\
(3)\qquad C &= \sum_u x_{iu}^4 = D + 2\alpha^4 \quad \text{for all} \quad i \\
D &= \sum_u x_{iu}^2 x_{ju}^2 = n_c \quad \text{for all} \quad i \neq j \\
& \qquad i, j = 1, \cdots, k; \qquad u = 1, \cdots, N
\end{aligned}
$$

Although this notation is not used in the previous work, we have found it convenient. All cross-product terms involving either x_i or x_ix_j are zero.

Box and Hunter [2] have shown that in order to have rotatability it is necessary that

$$
(4)\qquad C = 3D = D + 2\alpha^4
$$

so that $\alpha = n_c^{\frac{1}{4}}$. For a rotatable central composite design the variances and covariances of the coefficients are

$$
\begin{aligned}
V(b_0)/\sigma^2 &= (k+2)D/[(k+2)AD - kB^2] \\
V(b_i)/\sigma^2 &= 1/B \\
(5)\qquad V(b_{ii})/\sigma^2 &= [(k+1)AD - (k-1)B^2]/2D[(k+2)AD - kB^2] \\
V(b_{ij})/\sigma^2 &= 1/D \\
\operatorname{Cov}(b_0, b_{ii})/\sigma^2 &= -B/[(k+2)AD - kB^2] \\
\operatorname{Cov}(b_{ii}, b_{ij})/\sigma^2 &= (B^2 - AD)/2D[(k+2)AD - kB^2]
\end{aligned}
$$

The variance of a prediction $\hat{y}$ is given by

$$V(\hat{y})[(k+2)AD - kB^2]/\sigma^2 = (k+2)D + \rho^2(k+2)(AD - B^2)/B + \rho^4[(k+1)AD - (k-1)B^2]/2D \tag{6}$$

where $\rho^2 = \sum_{i=1}^{k} x_i^2$.

The quadratic coefficients will be correlated unless Cov $(b_{ii}, b_{ij}) = 0$. In order to have an orthogonal design, it therefore is necessary that $AD = B^2$, or that

$$n_0 = B^2/D - n_c - n_a \tag{7}$$

In determining the value of n_0 which would yield uniform information, Box and Hunter coded the variables so that $\sum_{u=1}^{N} x_{iu}^2 = N; i = 1, \cdots, k$. In these coded units the value of N which equates $V(\hat{y})$ at $\rho = 1$ to $V(\hat{y})$ at $\rho = 0$ is determined. To satisfy this requirement we multiply each x_{iu} by $(N/B)^{\frac{1}{2}} = c$. In these coded units we have $\sum_{u=1}^{N} x_{iu}^2 x_{ju}^2 = \lambda N$, $i < j$, so that $c^4 D = \lambda N$ and $N = B^2\lambda/D$. Substituting these values into (6) we find that uniform information is obtained when

$$\lambda = [k + 3 + \sqrt{9k^2 + 14k - 7}]/4(k+2) \tag{8}$$

Values of λ are given by Box and Hunter [2]. The number of center points needed to provide uniform information is

$$n_0 = B^2\lambda/D - n_c - n_a \tag{9}$$

In terms of the original coding (0, ± 1, and $\pm\alpha$), $V(\hat{y})$ at $\rho = 0$ equals $V(\hat{y})$ at $\rho = c$.

3. Partially Duplicated Second Order Response Surface Designs

In this paper eight types of partially duplicated second order response surface

Table 1. *Description of Designs*

Class	Type	n_c	n_a	Descriptions
0	A	2^k	$2k$	Full replication
	B	2^{k-1}	$2k$	Half replication
	C	2^{k-2}	$2k$	Quarter replication
1	I	$2^k + 2^k$	$2k$	Factorial duplicated
	II	$2^k + 2^{k-1}$	$2k$	Full replication plus half rep.
	III	$2^{k-1} + 2^{k-1}$	$2k$	Half replication, duplicated
	IV	$2^{k-1} + 2^{k-2}$	$2k$	Half plus quarter replication
	V	$2^{k-2} + 2^{k-2}$	$2k$	Quarter replication, duplicated
2	VI	2^k	$4k$	Full rep., star duplicated
	VII	2^{k-1}	$4k$	Half rep., star duplicated
	VIII	2^{k-2}	$4k$	Quarter rep., star duplicated

For all three classes of designs $D = n_c$. For classes 0 and 1 we have $A = D + 2k + n_0$, $B = D + 2\alpha^2$, and $\alpha^4 = D$. For class 2 we have $A = D + 4k + n_0$, $B = D + 4\alpha^2$, and $\alpha^4 = D/2$. Tables 2, 3, and 4 summarize some details of these designs.

designs are presented. These, in turn, may be broken into two classes of designs which, with the usual Box-Hunter designs indicated as class 0, are described in Table 1.

TABLE 2.

Values of n_c, n_a, and n_0 for Rotatable Designs With Full Factorials, $k = 2, \ldots, 8$ (O = orthogonal, U = uniform information)

Type		k = 2	3	4	5	6	7	8
I	n_c	8	16	32	64	128	256	512
	n_a	4	6	8	10	12	14	16
	$n_0(O)$	11	14	19	26	37	54	79
	$n_0(U)$	6	8	11	15	21	28	34
VI	n_c	4	8	16	32	64	128	256
	n_a	8	12	16	20	24	28	32
	$n_0(O)$	11	12	15	20	29	44	67
	$n_0(U)$	6	7	9	12	18	28	41
A	n_c	4	8	16	32	64	128	256
	n_a	4	6	8	10	12	14	16
	$n_0(O)$	8	9	12	17	24	35	52
	$n_0(U)$	5	6	7	10	15	21	28

TABLE 3.

Values of n_c, n_a, and n_0 for Rotatable Designs Where Half Replicates Are Possible $k = 5, \ldots, 8$ (O = orthogonal, U = uniform information)

Type		k = 5	6	7	8
II	n_c	48	96	192	384
	n_a	10	12	14	16
	$n_0(O)$	22	31	45	66
	$n_0(U)$	13	18	25	33
III	n_c	32	64	128	256
	n_a	10	12	14	16
	$n_0(O)$	17	24	35	52
	$n_0(U)$	10	15	21	28
VII	n_c	16	32	64	128
	n_a	20	24	28	32
	$n_0(O)$	11	16	25	40
	$n_0(U)$	6	9	16	25
B	n_c	16	32	64	128
	n_a	10	12	14	16
	$n_0(O)$	10	15	22	33
	$n_0(U)$	6	9	14	20

TABLE 4.

Values of n_c, n_a, and n_0 for Rotatable Designs Where Quarter Replicates are Possible $k = 8$ (O = orthogonal, U = uniform information)

Type		$k = 8$
IV	n_c	192
	n_a	16
	$n_0(O)$	43
	$n_0(U)$	25
V	n_c	128
	n_a	16
	$n_0(O)$	33
	$n_0(U)$	20
VIII	n_c	64
	n_a	32
	$n_0(O)$	21
	$n_0(U)$	13
C	n_c	64
	n_a	16
	$n_0(O)$	20
	$n_0(U)$	13

4. BLOCKING

As in the designs given by Box and Hunter, not all of the new designs can be orthogonally blocked and still retain rotatability. If it is desired to have orthogonal blocking, then a value of α can be found with a suitable choice of n_0 in each block such that the designs are nearly rotatable. Descriptions of some blocking arrangements are given for all three classes of designs in Table 5.

To achieve orthogonal blocking for classes 0 and 1, it is necessary to distribute the n_0 center points to the factorial and to the "star" designs so that

$$(D + n_{c0})/(2k + n_{a0}) = D/2\alpha^2 \tag{10}$$

where $n_{c0} + n_{a0} = n_0$. Since $n_{a0} = N - D - 2k - n_{c0}$, we can readily determine n_{c0} by

$$n_{c0} = N[D/(D + 2\sqrt{D})] - D \tag{11}$$

If, with a reasonable value of N, it is not possible to obtain a whole number for n_{c0}, then having chosen values for N and n_{c0}, the value for α may be determined by

$$\alpha^2 = (D/2)[(N - D - n_{c0})/(D + n_{c0})] \tag{12}$$

For class 2 the center points are to be distributed so that

$$(D + n_{c0})/(4k + n_{a0}) = D/4\alpha^2 \tag{13}$$

TABLE 5.

Orthogonal Blocking Arrangements for Some Central Composite Designs with Approximately Uniform Information

k = number of controlled variables
$n_c = D$ = number of points in cube (factorial portion)
n_{c0} = number of center points for blocks of cube
n_a = number of points in star ($2k$ for classes 0 and 1, $4k$ for class 2)
n_{a0} = number of center points for blocks of star
$N = n_c + n_{c0} + n_a + n_{a0}$ = total number of points
α = radius arm of star design

Class and Type	k	n_c	n_{c0}	n_a	n_{a0}	N	α*	Blocks**
0-A	2	4	2	4	2	12	1.414	4 + 2; 4+ 2
0-A	3	8	4	6	2	20	1.633(1.682)	2(4 + 2); 6 + 2
0-A	4	16	4	8	2	30	2.000	2(8 + 2); 8 + 2
0-A	5	32	8	10	4	54	2.366(2.378)	4(8 + 2); 10 + 4
0-B	5	16	6	10	1	33	2.000	16 + 6; 10 + 1
0-B	6	32	8	12	2	54	2.366(2.378)	2(16 + 4); 12 + 2
0-B	7	64	8	14	4	90	2.828	8(8 + 1); 14 + 4
0-C	8	64	12	16	3	95	2.828	4(16 + 3); 16 + 3
0-C	8	64	8	16	2	90	2.828	4(16 + 2); 16 + 2
1-I	2	8	2	4	3	17	1.673(1.682)	2(4 + 1); 4 + 3
1-I	3	16	4	6	4	30	2.000	4(4 + 1) or 2(8 + 2); 6 + 4
1-I	4	32	8	8	6	54	2.366(2.378)	4(8 + 2); 8 + 6
1-I	5	64	8	10	8	90	2.828	8(8 + 1) or 4(16 + 2); 10 + 8
1-II	5	48	6	10	6	70	2.667(2.632)	3(16 + 2); 10 + 6
1-III	5	32	8	10	4	54	2.366(2.378)	2(16 + 4); 10 + 4
1-III	5	32	4	10	3	49	2.404(2.378)	2(16 + 2); 10 + 3
1-III	6	64	8	12	6	90	2.828	4(16 + 2); 12 + 6
1-III	7	128	8	14	10	160	3.361(3.364)	8(16 + 1); 14 + 10
1-IV	8	192	12	16	13	233	3.694(3.722)	6(32 + 2); 16 + 13
1-IV	8	192	12	16	14	234	3.757(3.722)	6(32 + 2); 16 + 14
1-V	8	128	8	16	8	160	3.361(3.364)	8(16 + 1); 16 + 8
2-VI	2	4	3	8	2	17	1.195(1.189)	4 + 3; 2(4 + 1)
2-VI	3	8	6	12	2	28	1.414	2(4 + 3); 2(6 + 1)
2-VI	4	16	8	16	1	41	1.683(1.682)	2(8 + 4); 16 + 1
2-VI	4	16	10	16	2	44	1.664(1.682)	2(8 + 5); 2(8 + 1)
2-VI	5	32	10	20	1	63	2.000	2(16 + 5); 20 + 1
2-VI	5	32	12	20	2	66	2.000	4(8 + 3); 2(10 + 1)
2-VII	5	16	12	20	0	48	1.690(1.682)	16 + 12; 2(10 + 0)
2-VII	6	32	16	24	0	72	2.000	2(16 + 8); 2(12 + 0)
2-VII	7	64	16	28	0	108	2.366(2.378)	8(8 + 2); 2(14 + 0)
2-VIII	8	64	24	32	0	123	2.412(2.378)	4(16 + 6); 2(16 + 0)

* When the value of α for rotatability differs from the value of α for orthogonal blocking, the value for rotatability is given in parentheses.

** Given are the number of cube points in a block, the number of center points for the block containing the cube points; the number of points in the star, the number of center points for the block with the star points. Thus 2(4 + 2); 6 + 2 indicates three blocks, two of which contain 4 cube points and 2 center points and the third of which contains 6 star points and 2 center points.

Since $n_{a0} = N - D - 4k - n_{c0}$ we have

$$n_{c0} = N[D/(D + 4\sqrt{D/2})] - D \tag{14}$$

and

$$\alpha^2 = (D/4)[(N - D - n_{c0})/(D + n_{c0})] \tag{15}$$

Further blocking is possible for many of the partially duplicated designs. For class 2, type VI, $k = 4$, for example, we have $n_c = 2^4 = 16$, $n_a = 2(2k) = 16$, and $n_0 = 9$ for uniform information. The design is virtually rotatable if $n_{c0} = 8$ and $n_{a0} = 1$. Three blocks are possible:

1. A one-half replicate of the 2^4 plus 4 center points = 12,
2. The other half of the 2^4 plus 4 points = 12, and
3. The 8 star points, replicated, plus 1 center point = 17.

Another almost rotatable design is obtained with 12 center points by setting $n_{c0} = 10$ and $n_{a0} = 2$. Four blocks are possible here:

1. A one-half replicate of the 2^4 plus 5 center points = 13,
2. The other half of the 2^4 plus 5 center points = 13,
3. and 4. The star design plus 1 center point = 9 each.

5. Discussion

In this section the designs will be compared in certain important respects, namely, (1) biasses in the coefficients due to the presence of third order terms, (2) power of the test of the adequacy of the second order model (Lack of Fit Mean Square/Error Mean Square), and (3) information curves. In addition, the problem of testing for heterogeneous error is discussed. The following section will discuss some experimental strategies.

In order to compare the various designs, the designs must first be made equal in "size". Two designs will be considered equal in size if $\sum_u x_{iu}^2 = N$ for each design. This amounts to multiplying each x by $c = (N/B)^{\frac{1}{2}}$, if the values of x are 0, ± 1, and $\pm\alpha$. In coded units the mean squared distance of the experimental points from the center of the design is k. The moments as originally defined, $D = n_c = [iijj]$ for all $i \neq j$ and $B = [ii] = D + 2\alpha^2$ or $D + 4\alpha^2$ for all i, will be used in the following sub-sections.

a. *Bias Due to Third Order Terms*

The method described by Box and Hunter [2, page 198] is used. For all central composite designs covered in this paper it is found that the bias in the linear coefficients is given by

$$\mathcal{E}(b_i) = \beta_i + (ND/B^2)\left(\beta_{iii} + \sum_{j=1}^{k} \beta_{iij}\right). \tag{16}$$

Equation (16) shows that, for any n_c and n_a, the bias in the linear coefficients increases as n_0 increases. If the term ND/B^2 ($= \lambda$ for uniform information designs) is constant for the designs possible for a given k, then the biasses in the linear coefficients are equal for these designs.

For design types A, I, and VI none of the coefficients are biassed by the presence of β_{hij} terms, where $h \neq i \neq j$; $h, i, j, = 1, \cdots, k$. For $k \geq 6$ this also is true for design types B, II, III, and VII. For $k = 5$, however, design types B, II, III, and VII include a fractional factorial which has each of the two-factor interactions (β_{ij} terms) confounded with a three-factor interaction (β_{hlm} terms). Thus for $k = 5$,

$$\mathcal{E}(b_{ij}) = \beta_{ij} + z\beta_{hlm}, \tag{17}$$

where $z = (16/D)\,(N/B)^{\frac{1}{2}}$ for design types B, II, and VII and $z = (32/D)\,(N/B)^{\frac{1}{2}}$ for design type III. The bias is considerably less for design type II than for the other three design types ($z = 0.36$ vs. slightly larger than 1.).

b. *Testing for Lack of Fit*

The adequacy of the second order model is tested by comparing the Lack of Fit Mean Square with the Error Mean Square. For the types of central composite designs given in the paper the degrees of freedom for these mean squares are given in Table 6. With ND/B^2 constant as in the case of uniform

TABLE 6. *Degrees of Freedom for Lack of Fit and for Error*

Type of Design	Degrees of Freedom for Lack of Fit	Degrees of Freedom for Error
I	$2^k - k(k-1)/2$	$2^k + n_0 - 1$
II		$2^{k-1} + n_0 - 1$
VI		$2k + n_0 - 1$
A		$n_0 - 1$
III	$2^{k-1} - k(k-1)/2$	$2^{k-1} + n_0 - 1$
VII		$2k + n_0 - 1$
B		$n_0 - 1$
V	$2^{k-2} - k(k-1)/2$	$2^{k-2} + n_0 - 1$
VIII		$2k + n_0 - 1$
C		$n_0 - 1$
IV	[depends on the design]	

information designs, for a given k the new designs, compared with their Box and Hunter counterparts, have more degrees of freedom for error, so that the test for the adequacy of the second order model will have more power.

c. *Information Curves*

The information, $I(\hat{y})$, is defined as

$$I(\hat{y}) = \sigma^2/V(\hat{y}), \tag{18}$$

and the information at the center of the design, $I_0(\hat{y})$, is given by

$$I_0(\hat{y}) = \sigma^2/V(b_0) = n_0 + n_c + n_a - kB^2/(k+2)\,D \tag{19}$$

Certain values of $I_0(\hat{g})$ are given in Tables 7a and 7b. It is seen in these tables that only in a few designs the external points contribute substantially to the information at the center of the design.

TABLE 7a. *$I_0(\hat{g})$ for Design Types A, I, and VI for $k = 2, 3$, and 4*

Design Types	k 2	3	4
A	n_0	$n_0 + 0.01$	n_0
I	$n_0 + 0.34$	$n_0 + 0.40$	$n_0 + 0.92$
VI	$n_0 + 0.34$	$n_0 + 0.80$	$n_0 + 0.92$

TABLE 7b. *$I_0(\hat{g})$ for Design Types I, II, III, A, B, VI, and VII for $k = 5, 6$, and 7*

Design Type	k 5	6	7
I	$n_0 + 2.57$	$n_0 + 7.06$	$n_0 + 18.00$
II	$n_0 + 1.06$	$n_0 + 3.61$	$n_0 + 10.45$
III and A	$n_0 + 0.12$	$n_0 + 1.00$	$n_0 + 4.14$
B	$n_0 + 0.29$	$n_0 + 0.03$	$n_0 + 0.22$
VI	$n_0 + 0.57$	$n_0 + 0.06$	$n_0 + 0.44$
VII	$n_0 + 2.70$	$n_0 + 2.00$	$n_0 + 0.80$

Information curves are given by

$$I(\hat{g}) = \frac{2\lambda[(k+2)\lambda - k]N}{2(k+2)\lambda^2 + 2(k+2)\lambda(\lambda-1)\rho^2 + [(k+1)\lambda - (k-1)]\rho^4} \tag{20}$$

where $\lambda = ND/B^2$ and there are N observations. Since the value of λ depends on k, $I(\hat{g})$ is a function of k, N, and ρ. Comparing two uniform information designs with N_1 and N_2 runs, the information curves for every ρ will be in proportion N_1/N_2. Thus it is seen that any of the new designs, compared with the Box-Hunter designs, will give increased precision in direct proportion to the number of experimental runs.

d. *Testing for Heterogeneous Error*

In the new designs duplicates are obtained at the center of the design and either in the hypercube or the star design. Thus two estimates of error may be obtained, which we hypothesize are equal. The agreement between the two estimates may be tested by a two-tail F test.

It would seem that another method may be used if the duplicates are obtained in the hypercube. A factorial analysis of the absolute values of the differences between the pairs of duplicates should reveal the presence of heterogeneous error, if such exists.

6. Experimental Strategies

In experimental situations where the experimental error is fairly small, the designs of Class 2 will be found useful, since they permit partial duplication at minimum cost.

When little or nothing is known about the magnitudes of the effects relative to the experimental error, the designs of Class 1 might be used. A logical procedure would be to follow a sequential method of experimentation as suggested by C. Daniel [7], starting with a relatively small fractional factorial, except that center points are included in each block of runs. If no effects are found the experiment may be abandonded for a new experiment with the levels spread further, or additional blocks are run. Running additional blocks is continued at least until the factorial has two-factor interactions confounded with higher order interactions.

For $k = 5$, for example, we could begin with a one-half replication of the factorial, including some center points. Depending on the results (1) the star design and some center points can be run, to complete a type B design, (2) the block can be repeated and the star design and center points can be run, to complete a type III design, (3) the other half of the factorial can be run, again with some center points. If, after completing the factorial, still more information is needed, all or half of the factorial may be duplicated, again with center points. Eventually the star design will be run to complete the central composite design. The result will be a type I, type II, or type A design. In general we would prefer to use the Class 1 designs for this purpose, because the Class 2 designs require too many center points for orthogonal blocking, and also because, having started with the factorial, the radius arm of the star design for the Class 1 designs will have to be larger in order to achieve a rotatable composite design.

References

1. Box, G. E. P. (1954). The exploration and exploitation of response surfaces. Biometrics, 10, 16.
2. Box, G. E. P. and Hunter, J. S. (1957). Multifactor experimental designs for exploring response surfaces. Annals of Math. Stat., 28, 195.
3. Box, G. E. P. and Wilson, K. P. (1951). On the experimental attainment of optimum conditions. J. Royal Stat. Soc., Series B, 13, 1.
4. Davies, O. L., Editor (1954). Design and Analysis of Industrial Experiments, Oliver and Boyd, London.
5. DeBaun, R. M. (1956). Block effects in the determination of optimum conditions. Biometrics, 12, 20.
6. Dykstra, O. Jr. (1959). Partial duplication of factorial experiments. Technometrics, 1, 63.
7. Daniel, C. (1954). Fractional replication in industrial research. Proceedings of the Third Berkeley Symposium on Mathematical Statistics and Probability.

A Rank Sum Test for Comparing All Pairs of Treatments*

ROBERT G. D. STEEL

Mathematics Research Center, United States Army,

Madison, Wisconsin

A multiple comparison rank sum test, for the simultaneous comparison of all pairs of treatments in a one-way classification with equal numbers of observations, is presented. An example is worked and tables of critical values are given. Computation of probabilities for the general case of unequal numbers of observations is considered and means, variances, and covariances are given for this case.

1. INTRODUCTION

Recent research in multiple comparison procedures has been concerned with non-orthogonal sets of comparisons. Certainly some of the enthusiasm with which such problems have been approached is due to the natural desire to detect real differences among treatments even when this involves testing effects suggested by the data, for example, testing the difference between the greatest and least average treatment responses. Among such multiple comparison procedures, that which, in the framework of the analysis of variance, tests all possible pairs of treatment means, has received most attention.

In connection with multiple comparison procedures, several error rates have been defined. Two of these are the per comparison error rate and the experimentwise error rate. The former is defined as the ratio of the number of false conclusions drawn when the null hypothesis is true to the total number of conclusions drawn. This is the error rate customarily used when comparisons are independent, for example, in factorial experiments. An experimentwise error rate is defined as the proportion of experiments in which at least one false rejection of the null hypothesis is made.

Criticism has been levelled at both error rates. For a per comparison error rate, the probability of at least one false rejection of the null hypothesis increases with the number of treatments in the experiment. For an experimentwise error rate, the probability of detecting a real difference decreases as the number of treatments in the experiment increases. How applicable these facts are as criticisms depends upon such considerations as the size of meaningful differences, the importance of detecting them or failing to detect them or falsely detecting them, the number of treatments, and the choice of significance level for the chosen error rate.

*Sponsored by the United States Army under Contract No. DA-11-022-ORD-2059.

This paper presents a fixed rank sum test for comparing all pairs of treatments in a completely random design. Error rate is experimentwise so that the test is an analogue of Tukey's (3) fixed range test; this test is described by Federer (1) and others. For a fixed rank sum test, it is necessary to compute the probability distribution of a certain minimum rank sum. One can then test all pairs of treatments as in Section 3. The computation of probabilities is discussed in Section 4 and several distributions are presented in Tables 1 and 2. Means, variances and covariances are given for the general case.

Kruskal and Wallis (2) have given a rank test for one-way classifications which is an analogue of Snedecor's F-tests. Thus if it is concluded that the test criterion cannot reasonably be attributed to chance and the null hypothesis of no differences, one would probably conclude that at least the most extreme difference was evidence of a real difference between the corresponding population means. One could hardly conclude more without resorting to further testing, thus using a result-guided test procedure. The rank sum test presented here permits comparison of all treatment pairs without recourse to a result-guided procedure.

2. Procedure

Let $X_1 , \cdots , X_k$ be continuous random variables measuring a single characteristic on each of k treatments. Represent their cumulative distribution functions by $F_1 , \cdots , F_k$. We wish to test the null hypothesis

$$H_0 : F_1 = \cdots = F_k$$

using n observations from each population, against the alternative

$$H_1 : F_i \neq F_j , \quad \text{at least one } i, j \text{ pair.}$$

(More generally, one might take $n_1 , \cdots , n_k$ observations. However no tables have, as yet, been prepared for such a test.)

To test H_0 ,

1. Jointly rank the X_1's and X_2's, X_1's and X_3's, $\cdots$, X_1's and X_k's, X_2's and X_3's, $\cdots$, X_2's and X_k's, $\cdots$, and finally the X_{k-1}'s and X_k's, assigning rank 1 to the least observation and $2n$ to the greatest in each ranking.

2. Add the ranks of the X_1's when ranked with succeeding X_i's to give rank sums $T_{12} , \cdots , T_{1k}$; add the ranks of the X_2's when ranked with succeeding X_i's to give $T_{23} , \cdots , T_{2k}$; $\cdots$; finally $T_{k-1,k}$ is obtained. Compute conjugates $T'_{ij} = (2n + 1)n - T_{ij}$. The smaller of T_{ij} and T'_{ij} is to be used.

3. From Table 2 and using the predetermined significance level for k and n, extract the value of the significant rank sum. (For a one-tailed test, Table 1 gives significant rank sums. Such a test would be applied only if one were prepared to rank all treatments in stating H_1 ; if the treatments are then numbered according to their rank under H_1 , it is only necessary to compute the T_{1i}'s.)

4. Compare the tabulated rank sum with each of the computed rank sums, declaring any computed sum that is as small as or smaller than the tabulated sum to be significant. Error rate is experimentwise.

It may be noted that the procedure can be carried out by ordering the complete set of observations from smallest to largest and then ranking as described in step 1.

Let us investigate the effect of heterogeneity of variance in the parent population on the sensitivity of the proposed rank sum test as a test of location. Consider the extreme situation of two populations so different in variability that observations on one of them never lie within the range of observations on the other. In other words, assume one population to have variance which is virtually zero and the other to have a sizable variance. Then for n observations in each sample, the only possible values of the rank sum criterion are $n(n + 1)/2$ by steps of n to $n(3n + 1)/2$; the probabilities are given by the binomial distribution with $p = \frac{1}{2}$ and exponent n.

The most extreme sequence in this extreme situation occurs when all observations from the more variable population fall on the left (or right) of the observations from the less variable population. Let us compare the probability of this event with that for the same event and identical populations. The ratio is

$$\frac{1}{2^n}\frac{(2n)!}{(n!)^2} \approx \frac{1}{2^n}\frac{(2n+1)^{2n+\frac{1}{2}}e^{-(2n+1)}\sqrt{2\pi}}{(n+1)^{2n+1}e^{-2(n+1)}2\pi}$$

$$= \frac{1}{2^{n+\frac{1}{2}}}\frac{(2n)^{2n+\frac{1}{2}}(1+1/2n)^{2n+\frac{1}{2}}e}{n^{2n+1}(1+1/n)^{2n+1}\sqrt{\pi}}$$

using Stirling's approximation for factorials. Since $\lim_{n\to\infty}(1 + 1/n)^n = e$, this expression approaches $2^n/\sqrt{\pi n}$. Hence the ratio of these probabilities almost doubles itself for large n each time n increases by unity. For $n = 3, 4, 5$, exact values of the ratio are 2.5, 4.375 and 7.875 respectively.

It is unfortunate that the most extreme value of the test criterion is more frequent with heterogeneity of variance than with homogeneity. However, we have examined a most extreme case of heterogeneity and the most extreme term in this situation. In addition, we must bear in mind that means and variances are positively correlated in many experimental situations and a test which, in such a circumstance, detects differences between variances, even though they be attributed to differences between means, would be a useful test.

3. Example

The data used to illustrate the technique are a random sample from a larger set. The variable is air permeability of soil in units of square microns and measures the rapidity with which air flows through soil; low values are associated with compacted soils. All measurements were made on the same day. There is some evidence that the log normal distribution may be appropriate. There were three soil treatments.

For these data, the minimum value of the test criterion is 14, the conjugate of an observed rank total. From Table 2, we find that values less than or equal to 14 or equal to or greater than 22 occur with probability .625. There is no evidence of differences among treatments. (This conclusion does not hold for the complete set of data, 16 observations on each treatment.)

	Ranked with			Ranked with			Ranked with	
T_1	T_2	T_3	T_2	T_1	T_3	T_3	T_1	T_2
1.8	3	2	3.2	5	7	1.1	1	2
11.0	8	8	1.2	2	3	3.1	6	6
2.7	4	5	3.2	6	8	2.4	4	5
8.5	7	7	.7	1	1	2.2	3	4
Sums	22	22			19			
Criteria	14	14			17			

Data obtained through courtesy of Professor Champ B. Tanner, University of Wisconsin.

For the Kruskal-Wallis (3) procedure, all treatments are compared simultaneously but without benefit of multiple comparisons. Error rate may be described as experimentwise. The test criterion has the value $H = 1.885$, H being distributed approximately as χ^2 with $k - 1 = 2df$ when the null hypothesis is true. We find $.70 > P(H \geq 1.885) > .60$, essentially as found before. An important difference in the test criteria is that the former, a multiple comparison procedure, will point out which comparisons show significance.

If the Wilcoxon (4) two-sample test is used, then error rate is per comparison. The probability of obtaining a value of the test criterion less than or equal to 14, the observed minimum, or equal to or greater than 22, the complement of 14, is .342. Comparing this probability with an experimentwise error rate, we find that values less than or equal to 12 or equal to or greater than 24 occur with probability .260 while values less than or equal to 13 or equal to or greater than 23 occur with probability .417. In other words, the per comparison probability level of $P = .342$ for the minimum rank sum of 14, if chosen as an experimentwise error rate for the proposed multiple comparison test, would call for an observed minimum rank sum of 12. However, this value is more severe than required, P = .260, while 13 is not sufficiently severe, P = .417.

Since the only tie occurring in these data is for observations on the same treatment, no problem is introduced. When ties occur for observations in different treatments, it is suggested that they be assigned consecutive ranks in such a way as to increase or decrease the value of the test criterion, the decision depending upon the seriousness of the consequences of a wrong conclusion.

4. Mathematical Appendix: Distribution of min T_{ij}

Let $X_1, \cdots, X_k$ be continuous random variables measuring a single characteristic in k populations on which are made $n_1, \cdots, n_k$ observations respectively. There are $(\sum n_i)!/\pi(n_i!)$ possible permutations of such a set of observations, all permutations being equally likely under the null hypothesis. Hence if all possible permutations are considered, if ranks are assigned to the observations from the left by pairs of variables, and if the sum of the ranks associated with the variable of lower subscript is computed for each pair, then we generate

TABLE 1

(one-tailed—observed)

Probability that minimum T_{ij} is less than or equal to tabulated T_{ij} for $k = 3$.

$n = 1$		$n = 2$		$n = 3$		$n = 4$		$n = 5$		$n = 6$	
T_{ij}	P	T_{ij}	P	T_{ij}	P	T_{ij}	P	T_{ij}	P	T_{ij}	P
1	.8333										
2	1.0000										
		3	.3667								
		4	.6444								
		5	.9444								
		6	.9889	6	.1262						
		7	1.0000	7	.2381						
				8	.4333						
				9	.6649						
				10	.8363	10	.0388				
				11	.9399	11	.0753				
				12	.9869	12	.1440				
				13	.9976	13	.2389				
				14	.9994	14	.3800				
				15	1.0000	15	.5050	15	.0112		
						16	.6556	16	.0221		
						17	.7782	17	.0432		
						18	.8834	18	.0735		
						19	.9443	19	.1217		
						20	.9791				
						21	.9919			21	.0031
						22	.9981			22	.0062
						23	.9995			23	.0123
						24	.9999			24	.0216
						25	.9999				
						26	1.0000				

the distribution of $(T_{12}, \cdots, T_{1k}, T_{23}, \cdots, T_{2k}, \cdots, T_{k-1,k})$. The first subscript on any T_{ij} indicates the variable for which ranks are summed. From this joint distribution, we can obtain the distribution of $\min T_{ij}$.

Let $W(T_{12}, \cdots, T_{k-1,k} \mid n_1, \cdots, n_k)$ represent the number of ways of obtaining $(T_{12}, \cdots, T_{k-1,k})$ by summing the ranks of the X_1's, $\cdots$, X_{k-1}'s when ranked with each other and the X_k's as described in the preceding paragraph. If an X_1 is the last observation in a permutation, then the ranks $n_1 + n_2, \cdots, n_1 + n_k$ are in the rank sums $T_{12}, \cdots, T_{1k}$ respectively and in no others, and the remainder of each rank sum, namely $T_{1j} - n_1 - n_j$, must be obtained by drawing $n_1 - 1$ ranks from $n_1 + n_j - 1$ ranks, $j = 2, \cdots, k$. If an X_2 is last in a permutation, then the ranks $n_2 + n_3, \cdots, n_2 + n_k$ are in the rank sums $T_{23}, \cdots, T_{2k}$ respectively and in no others, and the remainder of each rank sum, namely $T_{2j} - n_2 - n_j$, must be obtained by drawing $n_2 - 1$ ranks from $n_2 + n_j - 1$ ranks, $j = 3, \cdots, k$. Finally, if an X_k is last in a permutation, no T_{ij} is affected. The number of ways in which $(T_{12}, \cdots, T_{k-1,k})$ may be

TABLE 2

(two-tailed)

Probability that either minimum observed T_{ij} or conjugate of maximum observed T_{ij} is less than or equal to T_{ij} for $k = 3$.

$n = 1$		$n = 2$		$n = 3$		$n = 4$		$n = 5$		$n = 6$	
T_{ij}	P	T_{ij}	P	T_{ij}	P	T_{ij}	P	T_{ij}	P	T_{ij}	P
1	1.0000										
		3	.6000								
		4	.9333								
		5	1.0000								
				6	.2286						
				7	.4143						
				8	.7000						
				9	.9393						
				10	1.0000	10	.0736				
						11	.1403				
						12	.2604				
						13	.4168				
						14	.6251				
						15	.7853	15	.0218		
						16	.9340	16	.0428		
						17	.9955	17	.0831		
						18	1.0000	18	.1405		
								19	.2298		
										21	.0062
										22	.0122
										23	.0240
										24	.0413

obtained is the sum of the numbers of ways in which each of these outcomes may occur. Thus we have equation 1.

$$
\begin{aligned}
&W(T_{12}, \cdots, T_{k-1,k} \mid n_1, \cdots, n_k) \\
&\quad = W(T_{12} - n_1 - n_2, \cdots, T_{1k} - n_1 - n_k, T_{23}, \cdots, T_{k-1,k} \\
&\qquad\qquad \mid n_1 - 1, n_2, \cdots, n_k) \\
&\quad + W(T_{12}, \cdots, T_{1k}, T_{23} - n_2 - n_3, \cdots, T_{2k} - n_2 - n_k, T_{34}, \cdots, T_{k-1,k} \\
&\qquad\qquad \mid n_1, n_2 - 1, n_3, \cdots, n_k) \qquad (1) \\
&\quad + \cdots \\
&\quad + W(T_{12}, \cdots, T_{k-2,k}, T_{k-1,k} - n_{k-1} - n_k \mid n_1, \cdots, n_{k-1} - 1, n_k) \\
&\quad + W(T_{12}, \cdots, T_{k-1,k} \mid n_1, \cdots, n_k - 1)
\end{aligned}
$$

Equation 1 may be used recursively till any term may be evaluated by inspection. Probabilities are obtained by dividing values given by equation 1 by the number of possible permutations.

For $k = 4$ and $n = 1$, it is readily demonstrable that equations 2, 3, and 4 are valid. It is then clear that they hold, where applicable, for all k and $n = 1$.

$$\mu(T_{ij}) = E(T_{ij}) = n_i(n_i + n_j + 1)/2 \tag{2}$$

$$\sigma^2(T_{ij}) = n_i n_j(n_i + n_j + 1)/12 \tag{3}$$

$$\left.\begin{aligned} &\sigma(T_{hi}T_{hj}) = n_h n_i n_j/12 = \sigma(T_{hj}T_{ij}) \\ &\sigma(T_{hi}T_{ij}) = -n_h n_i n_j/12 \\ &\sigma(T_{gh}T_{ij}) = 0 \end{aligned}\right\} \tag{4}$$

(Recall that the first of any pair of subscripts on a T is always the smaller.) It is only necessary to show that equations 2, 3, and 4 satisfy the obvious recursion formulas based on equation 1 in order to prove their general applicability. For convenience in handling subscripts, equations 2, 3, and 4 will be shown to be valid for $E(T_{12})$, $\sigma^2(T_{12})$, $\sigma(T_{12}T_{13})$, $\sigma(T_{12}T_{23})$, $\sigma(T_{13}T_{23})$ and $\sigma(T_{12}T_{34})$.

$$E(T_{12}) = \sum_{(T_{1s})} T_{12}W(T_{12}, \cdots, T_{k-1,k} \mid n_1, \cdots, n_k) \Big/ \sum_{(T_{1s})} W(T_{12}, \cdots, T_{k-1,k} \mid n_1, \cdots, n_k)$$

$$= \{\sum [(T_{12} - n_1 - n_2) + (n_1 + n_2)]W(T_{12} - n_1 - n_2, \cdots, T_{1k} - n_1 - n_k, T_{23}, \cdots, T_{k-1,k} \mid n_1 - 1, n_2, \cdots, n_k)$$

$$+ \sum T_{12}W(T_{12}, \cdots, T_{1k}, T_{23} - n_2 - n_3, \cdots, T_{2k} - n_2 - n_k, T_{34}, \cdots, T_{k-1,k} \mid n_1, n_2 - 1, n_3, \cdots, n_k)$$

$$+ \cdots$$

$$+ \sum T_{12}W(T_{12}, \cdots, T_{k-2,k}, T_{k-1,k} - n_{k-1} - n_k \mid n_1, \cdots, n_{k-1} - 1, n_k)$$

$$+ T_{12}W(T_{12}, \cdots, T_{k-1,k} \mid n_1, \cdots, n_{k-1}, n_k - 1)\} \Big/ \frac{(\sum n_i)!}{\pi(n_i!)}$$

$$= \left[\frac{(n_1 + 1)(n_1 + n_2)}{2} \cdot \frac{n_1}{\sum n_i} + \frac{n_1(n_1 + n_2)}{2} \cdot \frac{n_2}{\sum n_i}\right.$$

$$+ \frac{n_1(n_1 + n_2 + 1)}{2} \cdot \frac{n_3}{n_i} + \cdots + \frac{n_1(n_1 + n_2 + 1)}{2} \frac{n_{k-1}}{n_i}$$

$$\left. + \frac{n_1(n_1 + n_2 + 1)}{2} \frac{n_k}{n_i}\right] \Big/ \frac{(\sum n_i)!}{\pi(n_i!)} = \frac{n_1(n_1 + n_2 + 1)}{2}$$

Hence, equation 2 is valid.

The first term in the above expansion has been considered as two terms for purposes of summation. For n_1 and n_2, the set of T_{12}'s is $\sum_{i=1}^{n_1} i$ by steps of one to $\sum_{i=0}^{n_1-1} (n_1 + n_2 - i)$ and for $n_1 - 1$ and n_2, it is $\sum_{i=1}^{n_1-1} i$ by steps of one to $\sum_{i=1}^{n_1-1} (n_1 + n_2 - i)$. Hence in summing over the set of T_{12}'s, which is based on n_1 and n_2, the derived set of $(T_{12} - n_1 - n_2)$'s includes all the

values appropriate to the pair $(n_1 - 1, n_2)$, which appears in this first term, plus additional values which have zero probability in the distribution arising from $(n_1 - 1, n_2)$.

$$E(T_{12}^2) = \sum T_{12}^2 W(T_{12}, \cdots, T_{k-1,k} \mid n_1, \cdots, n_k) / \sum W(T_{12}, \cdots, T_{k-1,k} \mid n_1, \cdots, n_k)$$

$$= \{ \sum [(T_{12} - n_1 - n_2)^2 + 2(n_1 + n_2)(T_{12} - n_1 - n_2) + (n_1 + n_2)^2] W(T_{12} - n_1 - n_2, \cdots, T_{1k} - n_1 - n_k, T_{23}, \cdots, T_{k-1,k} \mid n_1 - 1, n_2, \cdots, n_k)$$

$$+ \sum T_{12}^2 W(T_{12}, \cdots, T_{1k}, T_{23} - n_2 - n_3, \cdots, T_{2k} - n_2 - n_k, T_{34}, \cdots, T_{k-1,k} \mid n_1, n_2 - 1, n_3, \cdots, n_k)$$

$$+ \cdots$$

$$+ \sum T_{12}^2 W(T_{12}, \cdots, T_{k-2,k}, T_{k-1,k} - n_{k-1} - n_k \mid n_1, \cdots, n_{k-1} - 1, n_k)$$

$$+ \sum T_{12}^2 W(T_{12}, \cdots, T_{k-1,k} \mid n_1, \cdots, n_{k-1}, n_k - 1) \} \Big/ \frac{(\sum n_i)!}{\pi n_i!}$$

$$= \frac{(n_1 - 1)n_2(n_1 + n_2)}{12} \frac{n_1}{\sum n_i} + \frac{n_1 - 1)^2(n_1 + n_2)^2}{4} \frac{n_1}{\sum n_i}$$

$$+ \frac{2(n_1 + n_2)n_1}{\sum n_i} \frac{(n_1 - 1)(n_1 + n_2)}{2} + \frac{n_1(n_1 + n_2)^2}{\sum n_i}$$

$$+ \left[\frac{n_1(n_2 - 1)(n_1 + n_2)}{12} + \frac{n_1^2(n_1 + n_2)^2}{4} \right] \frac{(\sum n_i - 1)!n_2}{\pi(n_i!)} \frac{\pi(n_i!)}{(\sum n_i)!}$$

$$+ \left[\frac{n_1 n_2(n_1 + n_2 + 1)}{12} + \frac{n_1^2(n_1 + n_2 + 1)^2}{4} \right] \frac{(\sum n_i - 1)!n_3}{\pi(n_i!)} \frac{\pi(n_i!)}{(\sum n_i)!}$$

$$+ \cdots$$

$$+ \frac{n_1 n_2(n_1 + n_2 + 1)}{12} \frac{n_{k-1}}{\sum n_i} + \frac{n_1^2(n_1 + n_2 + 1)^2}{4} \frac{n_{k-1}}{\sum n_i}$$

$$+ \frac{n_1 n_2(n_1 + n_2 + 1)}{12} \frac{n_k}{\sum n_i} + \frac{n_1^2(n_1 + n_2 + 1)^2}{4} \frac{n_k}{\sum n_i}$$

$$= \frac{n_1 n_2(n_1 + n_2 + 1)}{12} + \frac{n_1^2(n_1 + n_2 + 1)^2}{4}$$

Now,

$$\sigma^2(T_{12}) = E(T_{12}^2) - E^2(T_{12})$$

$$= \frac{n_1 n_2(n_1 + n_2 + 1)}{12}$$

and equation 3 is proven.

$$
\begin{aligned}
E(T_{12}T_{13}) = {} & \sum T_{12}T_{13}W(T_{12}, \cdots, T_{k-1,k} \\
& \qquad \mid n_1, \cdots, n_k) \Big/ \sum W(T_{12}, \cdots, T_{k-1,k} \mid n_1, \cdots, n_k) \\
= {} & \{ \sum [(T_{12} - n_1 - n_2)(T_{13} - n_1 - n_3) + (n_1 + n_2)(T_{13} - n_1 - n_3) \\
& + (n_1 + n_3)(T_{12} - n_1 - n_2) + (n_1 + n_2)(n_1 + n_3)] \\
& \qquad \cdot W(T_{12} - n_1 - n_2, \cdots, T_{k-1,k} \mid n_1 - 1, n_2, \cdots, n_k) \\
& + \sum T_{12}T_{13}W(T_{12}, \cdots, T_{1k}, T_{23} - n_2 - n_3, \cdots, T_{2k} - n_2 - n_k, \\
& \qquad T_{34}, \cdots, T_{k-1,k} \mid n_1, n_2 - 1, n_3, \cdots, n_k) \\
& + \cdots \\
& + \sum T_{12}T_{13}W(T_{12}, \cdots, T_{k-2,k}, T_{k-1,k} - n_{k-1} - n_k \\
& \qquad \mid n_1, \cdots, n_{k-1} - 1, n_k) \\
& + \sum T_{12}T_{13}W(T_{12}, \cdots, T_{k-1,k} \mid n_1, \cdots, n_{k-1}, n_k - 1) \} \Big/ \frac{(\sum n_i)!}{\pi(n_i!)} \\
= {} & \frac{(n_1 - 1)n_1n_2n_3}{12 \sum n_i} + \frac{(n_1 - 1)^2 n_1(n_1 + n_2)(n_1 + n_3)}{4 \sum n_i} \\
& + \frac{(n_1 - 1)n_1(n_1 + n_2)(n_1 + n_3)}{2 \sum n_i} + \frac{(n_1 - 1)n_1(n_1 + n_2)(n_1 + n_3)}{2 \sum n_i} \\
& + \frac{n_1(n_1 + n_2)(n_1 + n_3)}{\sum n_i} \\
& + \frac{n_1(n_2 - 1)n_2n_3}{12 \sum n_i} + \frac{n_1^2 n_2(n_1 + n_2)(n_1 + n_3 + 1)}{4 \sum n_i} \\
& + \frac{n_1n_2(n_3 - 1)n_3}{12 \sum n_i} + \frac{n_1^2 n_3(n_1 + n_2 + 1)(n_1 + n_3)}{4 \sum n_i} \\
& + \frac{n_1n_2n_3}{12} \frac{n_4}{\sum n_i} + \frac{n_1^2(n_1 + n_2 + 1)(n_1 + n_3 + 1)}{4} \frac{n_4}{\sum n_i} \\
& + \cdots \\
& + \frac{n_1n_2n_3}{12} \frac{n_k}{\sum n_i} + \frac{n_1^2(n_1 + n_2 + 1)(n_1 + n_3 + 1)}{4} \frac{n_k}{\sum n_i} \\
= {} & \frac{n_1n_2n_3}{12} + \frac{n_1^2(n_1 + n_2 + 1)(n_1 + n_3 + 1)}{4}
\end{aligned}
$$

Now

$$
\begin{aligned}
\sigma(T_{12}T_{13}) &= E(T_{12}T_{13}) - E(T_{12})E(T_{13}) \\
&= \frac{n_1n_2n_3}{12}
\end{aligned}
$$

which was to be shown.

$$E(T_{12}T_{23}) = \sum T_{12}T_{23}W(T_{12}, \cdots, T_{k-1,k} \mid n_1, \cdots, n_k) \Big/ \sum W(T_{12}, \cdots, T_{k-1,k} \mid n_1, \cdots, n_k)$$

$$= \{\sum [(T_{12} - n_1 - n_2) + (n_1 + n_2)]T_{23}W(T_{12} - n_1 - n_2, \cdots, T_{k-1,k} \mid n_1 - 1, n_2, \cdots, n_k)$$

$$+ \sum T_{12}T_{23}W(T_{12}, \cdots, T_{1k}, T_{23} - n_2 - n_3, \cdots, T_{2k} - n_2 - n_k, T_{34}, \cdots, T_{k-1,k} \mid n_1, n_2 - 1, n_3, \cdots, n_k)$$

$$+ \cdots$$

$$+ \sum T_{12}T_{13}W(T_{12}, \cdots, T_{k-1,k} \mid n_1, \cdots, n_{k-1}, n_k - 1)\} \Big/ \frac{(\sum n_i)!}{\pi(n_i!)}$$

$$= -\frac{(n_1 - 1)n_1n_2n_3}{12\sum n_i} + \frac{(n_1 - 1)n_1n_2(n_1 + n_2)(n_2 + n_3 + 1)}{4\sum n_i}$$

$$+ \frac{n_1n_2(n_1 + n_2)(n_2 + n_3 + 1)}{2\sum n_i}$$

$$- \frac{n_1(n_2 - 1)n_2n_3}{12\sum n_i} + \frac{n_1(n_2 - 1)n_2(n_1 + n_2)(n_2 + n_3)}{4\sum n_i}$$

$$+ \frac{n_1n_2(n_1 + n_2)(n_2 + n_3)}{2\sum n_i}$$

$$- \frac{n_1n_2(n_3 - 1)n_3}{12\sum n_i} + \frac{n_1n_2n_3(n_1 + n_2 + 1)(n_2 + n_3)}{4\sum n_i}$$

$$- \frac{n_1n_2n_3n_4}{12\sum n_i} + \frac{n_1n_2n_4(n_1 + n_2 + 1)(n_2 + n_3 + 1)}{4\sum n_i}$$

$$- \cdots$$

$$- \frac{n_1n_2n_3n_k}{12\sum n_i} + \frac{n_1n_2n_k(n_1 + n_2 + 1)(n_2 + n_3 + 1)}{4\sum n_i}$$

$$= -\frac{n_1n_2n_3}{12} + \frac{n_1n_2(n_1 + n_2 + 1)(n_2 + n_3 + 1)}{4}$$

Hence

$$\sigma(T_{12}T_{23}) = -\frac{n_1n_2n_3}{12}$$

which was to be shown.

The two remaining expected values are found similarly and equation 4 is seen to be valid.

Certain symmetry relations hold and are helpful in checking and computing tables required in conjunction with the test procedure. Consider a permutation of n_1 X_1's, $\cdots$, n_k X_k's. Ranks are assigned and the test criterion $(T_{12}, \cdots, T_{k-1,k})$ is computed. Corresponding to this permutation is one consisting of the same

sequence of X's in reverse order. When ranks are assigned, the test criterion becomes $(T'_{12}, \cdots, T'_{k-1,k})$ where $T'_{ij} = n_i(n_i + n_j + 1) - T_{ij}$ and both T_{ij} and T'_{ij} are the sums of n_i ranks, $n_i \leq n_j$. Hence, we have equation 5.

$$W(T_{12}, \cdots, T_{k-1,k} \mid n_1, \cdots, n_k) = W(T'_{12}, \cdots, T'_{k-1,k} \mid n_1, \cdots, n_k) \tag{5}$$

It is clear that for $n_1 = \cdots = n_k = n$, say, we may interchange subscripts and derive additional relations. This process together with application of equation 5 lead to equation 6 for $k = 3$. Clearly, such formulas can be generalized.

$$\left.\begin{aligned} W(T_{12}, T_{13}, T_{23}) &= W(T'_{12}, T'_{13}, T'_{23}) \\ &= W(T_{23}, T'_{12}, T'_{13}) = W(T'_{23}, T_{12}, T_{13}) \\ = W(T'_{13}, T'_{23}, T_{12}) &= W(T_{13}, T_{23}, T'_{12}) \\ &= W(T'_{23}, T'_{13}, T'_{12}) = W(T_{23}, T_{13}, T_{12}) \\ = W(T'_{12}, T_{23}, T_{13}) &= W(T_{12}, T'_{23}, T'_{13}) \\ &= W(T_{13}, T_{12}, T'_{23}) = W(T'_{13}, T'_{12}, T_{23}) \end{aligned}\right\} \tag{6}$$

In constructing the joint distribution, little use was actually made of equation 1 except for some checking. Instead, an enumeration procedure and application of equations 5 and 6 were found to be more amenable to what was basically a hand operation. Tables 1 and 2 were then obtained from the joint distribution by simple enumeration.

References

1. Federer, W. T. (1955). *Experimental Designs*, The Macmillan Company, New York.
2. Kruskal, W. H. and Wallis, W. A. (1952). Use of ranks in one-criterion variance analysis. *J. Amer. Stat. Assoc.*, 47, 583–621.
3. Tukey, J. W. (1953). *The problem of multiple comparisons*, unpublished memorandum in private circulation.
4. Wilcoxon, F. (1945). Individual comparisons by ranking methods. *Biometrics* 1, 80–83.

The Percentile Points of Distributions Having Known Cumulants

SIR RONALD A. FISHER
AND
E. A. CORNISH

In an earlier study of the uses of moments and cumulants in the specification of statistical distributions, the authors developed explicit asymptotic expansions, expressing any desired percentile point of such distributions in terms of known cumulants.

The general formulae are now presented as far as the sixth adjustment, based on the eighth cumulant, and also numerical tables showing the coefficients of all terms for ten chosen levels of significance over the range 0.5 to 0.0005 (single tail), together with the first five Hermite polynomials and tables for the common tests of significance, χ^2, t and z, at the same levels.

I. INTRODUCTORY

In 1937, in a study of the uses of moments and cumulants in the specification of statistical distributions, the authors (1937)* were led to develop explicit asymptotic expansions, expressing any desired percentile point of such distributions in terms of known cumulants. The general formulae were presented so far as the fourth adjustment, based on the sixth cumulant, and also numerical tables showing the coefficients of all terms for nine chosen levels of significance over the range 0.25 to 0.0005 (single tail), together with the first five Hermite polynomials at the same levels.

As an illustrative example, the cumulants of the z distribution were expressed in terms of $1/n_1$ and $1/n_2$, the reciprocals of the two numbers of degrees of freedom, and the rapid convergence at the 5% point exhibited for the case $n_1 = 24$ and $n_2 = 60$. In the intervening period, the formulae have frequently been found useful, either for calculations of higher accuracy in the case of functions already tabulated, or for values outside their range, or especially with tables of multiple entry to supply intermediate values more accurate than can be obtained by interpolation (Fisher 1941, Goldberg and Levine 1946), or for cases where no tables existed (Johnson and Welch 1939).

II. THE METHOD OF EXPANSION

The several steps in the method of expansion are set out in Sections 7 and 8 of our previous paper. Here a brief outline of these steps will be sufficient, as the principal purpose of this note is to extend the formulae and tables to the sixth corrective term, so widening the range of useful application.

If the element of frequency in the distribution of a variate ξ is $f(\xi)d\xi$, the

* Attention is drawn to this reference, as it has been repeatedly misquoted.

effect of the operator

$$\exp\left\{\frac{a_r}{r!}\left(-\frac{d}{d\xi}\right)^r\right\},$$

when acting on the frequency function, is to increase the rth cumulant by a_r, but leaves the distribution otherwise unchanged. This important operational property of the cumulants is basic to the method, the essential steps of which are as follows:

(i) If the cumulants $\kappa_1, \kappa_2, \cdots$ of the distribution of ξ are expressible in power series of the reciprocal of some number n, the frequency element may be represented as

$$\exp\left\{-av^{1/2}\frac{d}{d\xi}+\frac{1}{2}bv\frac{d^2}{d\xi^2}-\frac{1}{6}cv^{3/2}\frac{d^3}{d\xi^3}+\frac{1}{24}dv^2\frac{d^4}{d\xi^4}-\frac{1}{120}ev^{5/2}\frac{d^5}{d\xi^5}+\frac{1}{720}fv^3\frac{d^6}{d\xi^6}-\cdots\right\}\frac{1}{\sqrt{2\pi v}}e^{-(\xi-m)^2/2v}\,d\xi \tag{1}$$

where the coefficients a and c are of order $n^{-1/2}$, b and d of order n^{-1}, e of order $n^{-3/2}$, f of order $n^{-2}\cdots$, and are related, respectively, to κ_1 and κ_3, κ_2 and $\kappa_4, \kappa_5, \kappa_6, \cdots$, and m and v are the mean and variance of a normal distribution chosen for convenience.

(ii) Expanding the operator and integrating, the frequency less than $m + \xi v^{1/2}$ may be expressed in terms of the corresponding normal probability integral and a series of adjustments of decreasing order of magnitude.

(iii) If x is the normal deviate at some chosen level of probability, and ξ the corresponding deviate of the distribution under consideration, the difference $\xi - x$ may be found by equating the expression for the probability that the variate has a value less than ξ to

$$\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\xi}e^{-\xi^2/2}\,d\xi-\frac{1}{\sqrt{2\pi}}e^{-\xi^2/2}\left\{(\xi-x)-\frac{1}{2}(\xi-x)^2\xi+\frac{1}{6}(\xi-x)^3(\xi^2-1)-\frac{1}{24}(\xi-x)^4(\xi^3-3\xi)+\cdots\right\} \tag{2}$$

in which the coefficients are the Hermite polynomials. By considering the terms of each order of magnitude in succession, we may develop an expansion for $\xi - x$ in terms of successive polynomials in ξ.

(iv) The expansion for $\xi - x$ is converted to a much more useful expansion in terms of x, the values of which are known in advance, so obtaining the percentile deviate ξ explicitly in terms of the normal deviate.

The adjustments to the normal deviate having the required probability integral are set out in tabular form below. Adjustments V and VI are new with this paper.

Table I gives the numerical values of the first seven Hermite polynomials, over the same range of percentiles as given previously, and Table II gives the numerical values of the polynomials in the several adjustment terms.

TABLE I—*Hermite Polynomials*

p	x	$x^2 - 1$	$x^3 - 3x$	$x^4 - 6x^2 + 3$	$x^5 - 10x^3 + 15x$	$x^6 - 15x^4 + 45x^2 - 15$	$x^7 - 21x^5 + 105x^3 - 105x$
.5	0	−1.0000 0000 0000	0	3.0000 0000 0000	0	−15.0000 0000 0000	0
.25	0.6744 8975 0296	−0.5450 6357 6746	−1.7166 1929 6367	0.4773 4860 9677	7.1884 4393 0012	2.4617 8870 2987	−41.4702 1233 2512
.10	1.2815 5156 5545	0.6423 7441 5150	−1.7398 6719 3688	−4.1568 5277 1360	1.6322 4759 7878	22.8760 7332 1217	19.5233 8199 1054
.05	1.6448 5362 6951	1.7055 4345 4095	−0.4843 3791 7511	−5.9132 9534 2574	−7.7891 5362 1425	16.7544 5912 7787	74.2935 5459 2500
.025	1.9599 6398 4540	2.8414 5882 0694	1.6492 2898 3034	−5.2919 4705 3076	−16.9689 4156 4258	−6.7987 7905 6330	88.4882 8729 6296
.01	2.3263 4787 4041	4.4118 9443 1054	5.6109 0548 2094	−0.1827 6525 3449	−22.8687 9748 7187	−52.2869 5214 8942	15.5751 4495 1354
.005	2.5758 2930 3549	5.6348 9660 1021	9.3628 7318 0281	7.2124 7330 0116	−18.8733 9264 3621	−84.6770 0432 9402	−104.8731 5322 6683
.0025	2.8070 3376 8344	6.8794 3857 6622	13.6967 4885 5139	17.8089 2082 3031	−4.7967 5329 2545	−102.5092 6258 5743	−268.9664 1382 0596
.001	3.0902 3230 6168	8.5495 3570 6083	20.2395 8682 9338	36.8964 1796 5259	33.0001 5546 0763	−82.3185 2937 4518	−452.7443 1163 3935
.0005	3.2905 2673 1492	9.8275 6617 0663	25.7568 1572 7087	55.2707 9215 6104	78.8427 5615 2042	−16.9197 6407 7727	−528.7314 7290 0548

TABLE II—*Numerical values of the polynomials in the general formula*

	p									
	.5	.25	.10	.05	.025	.01	.005	.0025	.001	.0005
c	−0.16667	−0.09084	0.10706	0.28426	0.47358	0.73532	0.93915	1.14657	1.42492	1.63793
b	0	0.33724	0.64078	0.82243	0.97998	1.16317	1.28791	1.40352	1.54512	1.64526
d	0	−0.07153	−0.07249	−0.02018	0.06872	0.23379	0.39012	0.57070	0.84332	1.07320
c^2	0	0.07663	0.06106	−0.01878	−0.14607	−0.37634	−0.59171	−0.83890	−1.21026	−1.52234
bc	0.16667	0.09084	−0.10706	−0.28426	−0.47358	−0.73532	−0.93915	−1.14657	−1.42492	−1.63793
e	0.02500	0.00398	−0.03464	−0.04928	−0.04410	−0.00152	0.06010	0.14841	0.30747	0.46059
cd	−0.08333	0.00282	0.14644	0.17532	0.10210	−0.17621	−0.53531	−1.02868	−1.89358	−2.71243
c^3	0.05247	−0.01428	−0.11629	−0.11899	−0.02937	0.25195	0.59757	1.06301	1.86790	2.62337
b^2	0	−0.08431	−0.16019	−0.20561	−0.24500	−0.29079	−0.32198	−0.35088	−0.38628	−0.41132
bd	0	0.10729	0.10874	0.03027	−0.10308	−0.35068	−0.58518	−0.85605	−1.26497	−1.60980
bc^2	0	−0.19158	−0.15265	−0.04696	0.36517	0.94084	1.47928	2.09726	3.02565	3.80584
f	0	0.00998	0.00227	−0.01082	−0.02357	−0.03176	−0.02621	−0.00666	0.04592	0.10950
ce	0	−0.05126	0.01086	0.09462	0.16106	0.16058	0.05366	−0.17498	−0.70466	−1.30531
d^2	0	−0.03285	0.00776	0.05985	0.09659	0.07888	−0.01226	−0.19116	−0.59062	−1.03555
c^2d	0	0.14764	−0.10858	−0.39517	−0.55856	−0.32621	0.35696	1.60445	4.29316	7.23307
c^4	0	−0.06898	0.09585	0.25623	0.31624	0.07286	−0.46535	−1.39199	−3.32716	−5.40702
b^2c	−0.16667	−0.09084	0.10706	0.28426	0.47358	0.73532	0.93915	1.14657	1.42492	1.63793
be	−0.05000	−0.00796	0.06928	0.09855	0.08820	0.00305	−0.12021	−0.29682	−0.61494	−0.92118
bcd	0.25000	−0.00846	−0.43931	−0.52597	−0.30631	0.52864	1.60592	3.08604	5.68074	8.13729
bc^3	−0.20988	0.05714	0.46515	0.47598	0.11748	−1.00781	−2.39028	−4.25205	−7.47158	−10.49350
g	−0.00298	0.00049	0.00454	0.00332	−0.00135	−0.01037	−0.01680	−0.02034	−0.01633	−0.00336
cf	0.02083	−0.00791	−0.03371	−0.01141	0.04024	0.12188	0.16262	0.15484	0.01974	−0.21672
de	0.03333	−0.01168	−0.04871	−0.01011	0.07080	0.18338	0.21571	0.14718	−0.18946	−0.70228
c^2e	−0.08333	0.04892	0.12788	−0.03219	−0.31878	−0.67431	−0.71959	−0.37377	1.00984	3.01445
cd^2	−0.11111	0.05854	0.14979	−0.06496	−0.42767	−0.82209	−0.76354	−0.11687	2.10109	5.17256
c^3d	0.23457	−0.17412	−0.28497	0.38896	1.39396	2.27932	1.73496	−0.69355	−8.16706	−18.12350
c^5	−0.09091	0.08374	0.07657	−0.28485	−0.76311	−1.05359	−0.53986	1.05630	5.56273	11.36150

cd^2	−0.11111	0.05854	0.14979	−0.06496	−0.42767	−0.82209	−0.76354	−0.11687	2.10169	5.17250
c^2d	0.23457	−0.17412	−0.28497	0.38896	1.39396	2.27932	1.73496	−0.69355	−8.16706	−18.12350
c^5	−0.09091	0.08374	0.07657	−0.28485	−0.76311	−1.05359	−0.53986	1.05630	5.56273	11.36150

TABLE II—*Continued*

	p									
	.5	.25	.10	.05	.025	.01	.005	.0025	.001	.0005
b^3	0	0.04216	0.08010	0.10280	0.12250	0.14540	0.16099	0.17544	0.19314	0.20566
b^2d	0	−1.13411	−0.13593	−0.03784	0.12885	0.43835	0.73147	1.07006	1.58122	2.01225
b^2c^2	0	0.33526	0.26714	−0.08217	−0.63904	−1.64648	−2.58873	−3.67020	−5.29489	−6.66023
bf	0	−0.02496	−0.00567	0.02705	0.05892	0.07941	0.06553	0.01666	−0.11479	−0.27376
bce	0	0.17941	−0.03802	−0.33116	−0.56370	−0.56204	−0.18780	0.61244	2.46632	4.56859
bd^2	0	0.11498	−0.02716	−0.20949	−0.33807	−0.27606	0.04292	0.66905	2.06717	3.62442
bc^2d	0	−0.66436	0.48861	1.77828	2.51350	1.46795	−1.60634	−7.22002	−19.31920	−32.54883
bc^4	0	0.37938	−0.52718	−1.40927	−1.73933	−0.40075	2.55940	7.65594	18.29937	29.73862
h	0	−0.00103	0.00048	0.00184	0.00219	0.00039	−0.00260	−0.00642	−0.01123	−0.01311
cg	0	0.00933	−0.00937	−0.02175	−0.01828	0.01651	0.06034	0.10752	0.14686	0.13160
df	0	0.01700	−0.01657	−0.03576	−0.02362	0.04640	0.12406	0.19389	0.20794	0.09578
e^2	0	0.01023	−0.01020	−0.02156	−0.01387	0.02788	0.07209	0.10787	0.09883	0.00803
c^2f	0	−0.04600	0.07442	0.12464	0.04811	−0.25595	−0.56524	−0.81931	−0.79570	−0.24501
cde	0	−0.14366	0.22074	0.34077	0.05220	−0.91516	−1.77648	−2.26940	−1.31805	1.58689
d^3	0	−0.03120	0.04431	0.06346	−0.00731	−0.21997	−0.39141	−0.45485	−0.12354	0.67346
c^3e	0	0.15345	−0.35265	−0.42236	0.15607	1.76170	3.04573	3.56562	1.33265	−4.18956
c^2d^2	0	0.30132	−0.63747	−0.68830	0.52403	3.60694	5.77533	5.99761	−0.47341	−13.97518
c^4d	0	−0.36979	1.09609	0.79893	−1.77731	−7.45279	−10.76740	−9.59341	6.35417	36.18663
c^6	0	0.10100	−0.40194	−0.12422	1.07691	3.38425	4.37679	3.02544	−5.79725	−20.92219

III. Algebraic Forms of the Expansion for Common Tests of Significance

In general applications of the formulae given above, the coefficients $a, b, c, \cdots$ can be evaluated numerically, and their values substituted, but to facilitate the use of the formulae with those statistics in frequent use, the substitution has been made algebraically to give the general form of the expansion in terms of the parameters of the distributions concerned.

Adjustment	Coefficient	Polynomial in x	Divisor
I	a	1	1
	c	$x^2 - 1$	6
II	b	x	2
	d	$x^3 - 3x$	24
	c^2	$-(2x^3 - 5x)$	36
III	bc	$-(x^2 - 1)$	6
	e	$x^4 - 6x^2 + 3$	120
	cd	$-(x^4 - 5x^2 + 2)$	24
	c^3	$12x^4 - 53x^2 + 17$	324
IV	b^2	$-x$	8
	bd	$-(x^3 - 3x)$	16
	bc^2	$5(2x^3 - 5x)$	72
	f	$x^5 - 10x^3 + 15x$	720
	ce	$-(2x^5 - 17x^3 + 21x)$	180
	d^2	$-(3x^5 - 24x^3 + 29x)$	384
	c^2d	$14x^5 - 103x^3 + 107x$	288
	c^4	$-(252x^5 - 1688x + 1511x)$	7776
V	b^2c	$x^2 - 1$	6
	be	$-(x^4 - 6x^2 + 3)$	60
	bcd	$x^4 - 5x^2 + 2$	8
	bc^3	$-(12x^4 - 53x^2 + 17)$	81
	g	$x^6 - 15x^4 + 45x^2 - 15$	5040
	cf	$-(x^6 - 13x^4 + 33x^2 - 9)$	432
	de	$-(x^6 - 12x^4 + 29x^2 - 8)$	240
	c^2e	$16x^6 - 181x^4 + 393x^2 - 90$	1080
	cd^2	$12x^6 - 129x^4 + 271x^2 - 64$	576
	c^3d	$-(80x^6 - 803x^4 + 1513x^2 - 304)$	1296
	c^5	$960x^6 - 8937x^4 + 15062x^2 - 2651$	29160

Adjustment	Coefficient	Polynomial in x	Divisor
VI	b^3	x	16
	b^2d	$5(x^3 - 3x)$	64
	b^2c^2	$-35(2x^3 - 5x)$	288
	bf	$-(x^5 - 10x^3 + 15x)$	288
	bce	$7(2x^5 - 17x^3 + 21x)$	360
	bd^2	$7(3x^5 - 24x^3 + 29x)$	768
	bc^2d	$-(14x^5 - 103x^3 + 107x)$	64
	bc^4	$11(252x^5 - 1688x^3 + 1511x)$	15552
	h	$x^7 - 21x^5 + 105x^3 - 105x$	40320
	cg	$-(2x^7 - 37x^5 + 160x^3 - 135x)$	5040
	df	$-(x^7 - 17x^5 + 69x^3 - 57x)$	1152
	e^2	$-(2x^7 - 33x^5 + 132x^3 - 108x)$	3600
	c^2f	$18x^7 - 293x^5 + 1100x^3 - 795x$	5184
	cde	$18x^7 - 273x^5 + 974x^3 - 695x$	1440
	d^3	$9x^7 - 131x^5 + 451x^3 - 321x$	3072
	c^3e	$-(396x^7 - 5708x^5 + 18755x^3 - 11811x)$	19440
	c^2d^2	$-(594x^7 - 8193x^5 + 26006x^3 - 16367x)$	13824
	c^4d	$5148x^7 - 67004x^5 + 195259x^3 - 109553x$	62208
	c^6	$-(154440x^7 - 1887684x^5 + 5033714x^3 - 2542637x)$	4199040

(a) χ^2 *distribution*

If n is the number of degrees of freedom

$$\begin{aligned}\chi^2 &= n \\ &+ \sqrt{n}\,(x\sqrt{2}) \\ &+ \tfrac{2}{3}(x^2 - 1) \\ &+ \frac{1}{\sqrt{n}}\left(\frac{x^3 - 7x}{9\sqrt{2}}\right) \\ &- \frac{1}{n}\left(\frac{6x^4 + 14x^2 - 32}{405}\right) \\ &+ \frac{1}{n\sqrt{n}}\left(\frac{9x^5 + 256x^3 - 433x}{4860\sqrt{2}}\right) \\ &+ \frac{1}{n^2}\left(\frac{12x^6 - 243x^4 - 923x^2 + 1472}{25515}\right) \\ &- \frac{1}{n^2\sqrt{n}}\left(\frac{3753x^7 + 4353x^5 - 289517x^3 - 289717x}{9185400\sqrt{2}}\right).\end{aligned} \tag{3a}$$

(b) *t distribution*

If n is the number of degrees of freedom

$$\begin{aligned} t = {} & x \\ & + \frac{1}{n}\left(\frac{x^3 + x}{4}\right) \\ & + \frac{1}{n^2}\left(\frac{5x^5 + 16x^3 + 3x}{96}\right) \\ & + \frac{1}{n^3}\left(\frac{3x^7 + 19x^5 + 17x^3 - 15x}{384}\right) \\ & + \frac{1}{n^4}\left(\frac{79x^9 + 776x^7 + 1482x^5 - 1920x^3 - 945x}{92160}\right) \\ & + \frac{1}{n^5}\left(\frac{27x^{11} + 339x^9 + 930x^7 - 1782x^5 - 765x^3 + 17955x}{368640}\right). \end{aligned} \tag{3b}$$

The number of terms and the orders of magnitude differ in this expansion because it has been derived from the original expansion of Student's integral in powers of n^{-1} as given by Fisher (1926).

(c) *z distribution*

If n_1 and n_2 are the degrees of freedom, and

$$\sigma = \frac{1}{n_1} + \frac{1}{n_2} \quad \delta = \frac{1}{n_1} - \frac{1}{n_2},$$

then

$$\begin{aligned} z = {} & \sqrt{\frac{\sigma}{2}}\,(x) \\ & - \delta\left(\frac{x^2 + 2}{6}\right) \\ & + \sqrt{\frac{\sigma}{2}}\left\{\sigma\left(\frac{x^3 + 3x}{24}\right) + \frac{\delta^2}{\sigma}\left(\frac{x^3 + 11x}{72}\right)\right\} \\ & - \left\{\delta\sigma\left(\frac{x^4 + 9x^2 + 8}{120}\right) - \frac{\delta^3}{\sigma}\left(\frac{3x^4 + 7x^2 - 16}{3240}\right)\right\} \\ & + \sqrt{\frac{\sigma}{2}}\left\{\sigma^2\left(\frac{x^5 + 20x^3 + 15x}{1920}\right) + \delta^2\left(\frac{x^5 + 44x^3 + 183x}{2880}\right)\right. \\ & \left. + \frac{\delta^4}{\sigma^2}\left(\frac{9x^5 - 284x^3 - 1513x}{155520}\right)\right\} + \left\{\delta\sigma^2\left(\frac{4x^6 - 25x^4 - 177x^2 + 192}{20160}\right)\right. \\ & \left. + \delta^3\left(\frac{4x^6 + 101x^4 + 117x^2 - 480}{90720}\right) - \frac{\delta^5}{\sigma^2}\left(\frac{12x^6 + 513x^4 + 841x^2 - 2560}{1632960}\right)\right\} \end{aligned} \tag{3c}$$

$$- \sqrt{\frac{\sigma}{2}}\left\{\sigma^3\left(\frac{x^7 + 7x^5 + 7x^3 + 105x}{21504}\right)\right.$$

$$+ \delta\sigma^2\left(\frac{801x^7 + 10511x^5 + 30151x^3 + 62241x}{4838400}\right)$$

$$- \frac{\delta^4}{\sigma}\left(\frac{477x^7 + 4507x^5 - 82933x^3 - 264363x}{43545600}\right)$$

$$\left. + \frac{\delta^6}{\sigma^3}\left(\frac{3753x^7 + 55383x^5 - 368897x^3 - 1213927x}{1175731200}\right)\right\}.$$

As an illustration of the accuracy, we may use the example given previously. When $n_1 = 24$ and $n_2 = 60$, the 5% value of z is 0.26534844, and the asymptotic expansion yields the following values:

Order of magnitude	Successive terms	Successive totals	Successive errors
0	0.2809 1224	0.2809 1224	0.0155 6380
1	— 196 0643	2613 0581	— 40 4263
2	44 6851	2657 7432	4 2588
3	— 4 8004	2652 9428	— 5416
4	5645	2653 5073	229
5	— 154	2653 4919	75
6	— 102	2653 4817	— 27

The numerical values of the polynomials in x occurring in the above formulae are given in Table III.

IV. Examples of the Types of Problem to Which the Expansion Has Been Applied

(a) *The asymptotic approach to Behrens' integral*

Fisher (1926) developed the ordinate and integral of Student's distribution in a series of powers of n^{-1}, giving the polynomial coefficients so far as the fifth adjustment. The purpose of this expansion was to supply sufficiently accurate values of the probabilities corresponding to any values of t for values of n beyond the range which it was proposed to tabulate.

With Behrens' extension of Student's test there were even stronger reasons for using a similar method. The direct calculations carried out by Sukhatme (1938) are very much more laborious than those needed for Student's integral. At any single level of significance, various values are required for three parameters, provided by the two numbers of degrees of freedom of the two samples, and the estimated ratio of the variances of the two means. For functions of many variables, there is a great advantage in the use of explicit formulae in which the several variables may be substituted, and there is much to be gained by extending the use of such formulae over regions too extensive for complete tabulation. Finally, it should be noted that the logical situation in which we would prefer to rely on the separate estimates of variances from the two samples rather than on any process of pooling these estimates, is of more frequent occurrence with large samples than with small, and is particularly applicable to cases,

TABLE III—*Numerical values of the polynomials in the expansions for χ^2, t and z.*

(a) χ^2

p									
.5	.25	.10	.05	.025	.01	.005	.0025	.001	.0005
0	0.9538726	1.8123876	2.3261743	2.7718076	3.2899527	3.6427727	3.9697452	4.3702484	4.6535075
−0.666667	−0.363376	0.428250	1.137029	1.894306	2.941263	3.756598	4.586292	5.699690	6.551711
0	−0.346842	−0.539450	−0.554981	−0.486382	−0.290266	−0.073888	0.193953	0.619006	0.989534
0.07901	0.06022	−0.01772	−0.12296	−0.27240	−0.54197	−0.80252	−1.11315	−1.60211	−2.03211
0	−0.0309	0.0022	0.0779	0.1948	0.4116	0.6228	0.8752	1.2735	1.6249
0.0577	0.0393	−0.0253	−0.1006	−0.1952	−0.3425	−0.4642	−0.5886	−0.7467	−0.8535
0	0.012	0.073	0.122	0.170	0.203	0.183	0.100	−0.145	−0.469

The values in the columns are in the same order as the polynomials in expansion 3a. Sufficient figures have been retained to ensure accuracy in the fourth decimal place for $n > 30$, except for $x\sqrt{2}$ which should be taken more accurately for $n > 1600$.

(b) t

p									
.5	.25	.10	.05	.025	.01	.005	.0025	.001	.0005
0	0.24533	0.84658	1.52377	2.37227	3.72907	4.91655	6.23122	8.15013	9.72973
0	0.0795	0.5709	1.4202	2.8225	5.7197	8.8348	12.8509	19.6925	26.1330
0	−0.005	0.259	0.983	2.556	6.719	12.144	20.221	36.154	53.169
0	0	0.1	0.4	1.6	5.6	12.1	23.2	48.6	79.4

The values in the columns are in the same order as the polynomials in expansion 3b. Sufficient figures have been retained to ensure accuracy in the fourth decimal place for $n > 30$.

The values in the columns are in the same order as the polynomials in expansion 3b. Sufficient figures have been retained to ensure accuracy in the fourth decimal place for $n > 30$.

TABLE III—*Continued*

(c) z

p									
.5	.25	.10	.05	.025	.01	.005	.0025	.001	.0005
0	0.67448975	1.28155157	1.64485363	1.95996398	2.32634787	2.57582930	2.80703377	3.09023231	3.29052673
0.33333333	0.40915607	0.60706240	0.78425724	0.97357647	1.23531574	1.43914943	1.64657310	1.92492262	2.13792770
0	0.0970966	0.2478934	0.3910327	0.5587089	0.8153747	1.0340770	1.2724563	1.6158742	1.8958323
0	0.1073089	0.2250258	0.3131057	0.4040101	0.5302747	0.6308956	0.7360447	0.8819839	0.9975582
0.0666667	0.1025116	0.2123230	0.3305821	0.4777495	0.7166304	0.9311327	1.1750042	1.5428288	1.8557024
−0.004938	−0.003764	0.001108	0.007685	0.017025	0.033873	0.050157	0.069572	0.100132	0.127007
0	0.008539	0.033737	0.065478	0.108805	0.184807	0.257207	0.343093	0.478317	0.597758
0	0.047595	0.114789	0.176687	0.249610	0.363825	0.464148	0.576788	0.745061	0.887356
0	−0.00711	−0.01611	−0.02343	−0.03114	−0.04168	−0.04971	−0.05761	−0.06765	−0.07475
0.00952	0.00529	−0.00736	−0.01938	−0.03126	−0.04286	−0.04537	−0.03958	−0.01462	0.02094
−0.00529	−0.00447	0	0.00722	0.01859	0.04128	0.06515	0.09556	0.14695	0.19516
0	0	0	0	0	0	0.0178	0.0256	0.0384	0.0502
0	0.00344	0.00833	0.01491	0.02660	0.05478	0.09004	0.14149	0.24158	0.34748
0	0.0109	0.0381	0.0804	0.1534	0.3174	0.5105	0.7799	1.2814	1.7939
0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0

The values in the columns are in the same order as the polynomials in expansion 3c. Sufficient figures have been retained to ensure accuracy in the sixth decimal place for $n_1 > 24$ and $n_2 > 60$.

such as arise in Physics and Astronomy, in which we wish to compare estimates of the value of the same quantity (a) from relatively ample data of low intrinsic accuracy, and (b) from a small series of observations of relatively high precision. When, as often happens, the estimates of precision of the means obtained in these two ways are of the same order of magnitude, the only satisfactory test is that based on Behrens' solution. The asymptotic expansion is particularly suitable for evaluating the percentiles for this special application. There were thus four manifest advantages of the asymptotic approach to Behrens' integral (Fisher 1941):

(i) a check on Sukhatme's values, obtained by a completely independent method, and applicable at least for the higher values of n_1 and n_2 ,

(ii) greater accuracy than could be obtained for percentiles from Sukhatme's table for values of n_1 and n_2 greater than 12,

(iii) a wider range of levels of significance in the region to which the asymptotic expansion is applicable,

(iv) the theoretical guidance offered by the algebraic form of the leading terms of the expansion.

(b) *The fiducial distribution of the binomial parameter, p*

When discussing the application of the fiducial argument to discontinuous observations, Fisher (1959) found that the mean of the fiducial distribution of p, the parameter of the binomial distribution, for given observational frequencies a, b out of N, was

$$\bar{p} = \frac{a}{N} + \frac{b-a}{2N^2} - \frac{3(b-a)}{2N^3} + \frac{15(b-a)}{2N^4} - \cdots \tag{4}$$

if χ were taken to be normally distributed.

On the other hand, the mean of the Bayesian distribution *a posteriori*, using the Bayesian probability *a priori*

$$\frac{1}{\pi\sqrt{pq}}\,dp,$$

was

$$\frac{a+\frac{1}{2}}{N+1} = \frac{a}{N} + \frac{b-a}{2N^2} - \frac{b-a}{2N^3} + \frac{b-a}{2N^4} - \cdots . \tag{5}$$

Asymptotic agreement between these means appears when allowance is made for the effects of departure from normality in the binomial distribution, which are appreciable in the expression (4).

Direct application of the asymptotic expansion, using the six adjustment terms gave the following expansion for the binomial variate in terms of the normal deviate x

$$a = pN + x\sqrt{Npq} + \tfrac{1}{6}(q-p)(x^2-1) + \frac{1}{72\sqrt{Npq}}\{-x^3 + x - pq(2x^3 - 14x)\} + \cdots$$

which, after inversion, gave*

$$p = \frac{a}{N} - \frac{x\sqrt{ab}}{N^{3/2}} + \frac{(b-a)(2x^2+1)}{6N^{2}} + \left\{\frac{(-2N^2+26ab)x^3 + (-7N^2+34ab)x}{72N^{5/2}\sqrt{ab}}\right\} + \cdots \tag{6}$$

whence, substituting its average value for each power of x, the mean of the fiducial distribution is

$$\bar{p} = \frac{a}{N} + \frac{b-a}{2N^2} - \frac{b-a}{2N^3} + \frac{b-a}{2N^4} - \cdots \tag{7}$$

agreeing so far as the fourth term with (5).

The expansion (6) also provides a ready means for comparing the two distributions with respect to other properties. For example, although the means are in agreement, the asymptotic fiducial distribution has the higher variance.

See also the alternative treatment in Fisher (1957).

(c) *Quantitative inheritance*

Panse (1940) has described a statistical technique for the study of quantitative inheritance, in which genetic models, based on data from the F_2 and F_3 generations, are used to represent the constitution of particular characters. The statistical consequences in the population, corresponding to these models were assessed, using the cumulant function of the joint distribution of the F_2 phenotypic value, the mean of the F_3 progeny, and the genotypic variance of F_3 progeny. These functions provided the data for expressing an attribute of the F_3 progeny in terms of the F_2 phenotypic values, and thus the effects of selection in the F_2 phenotype on the mean value of the F_3 progeny could be determined by integration over the F_2 distribution. When the intensity of selection was assigned, the limits of integration were calculable from the asymptotic expansion of the deviate.

V. A Class of Distributions with a Finite Condensation at Zero

The Poisson Series, a discontinuous distribution of positive integers, is well known to have the simple series of cumulants

$$\kappa_r = m,$$

for all values of r. Correspondingly, the cumulative function is

$$K = m(e^{it} - 1),$$

and the characteristic function

$$M = \exp\{m(e^{it} - 1)\}.$$

* For the remaining terms see Fisher (1959)

It is less well known that if, for all values of r,

$$\kappa_r = r!m,$$

or, if a scaling factor be introduced,

$$\kappa_r = r!a^r m,$$

the distribution, derived from

$$K = \frac{mait}{1 - ait}$$

and

$$M = e^{-m}e^{m/(1-ait)}$$

$$= e^{-m} \sum_0^\infty \frac{m^n}{n!}(1 - ait)^{-n}$$

is continuous over the range of positive values, with a finite condensation at zero. For $(1 - ait)^{-n}$ is the characteristic function of the Eulerian distribution

$$\frac{1}{(n-1)!}\left(\frac{X}{a}\right)^{n-1} e^{-x/a}\frac{dX}{a}$$

or of

$$\chi^2 = \frac{2X}{a},$$

for $2n$ degrees of freedom, or of the sum of n random variables, each distributed as

$$e^{-x/a}\frac{dx}{a}, \; x \geq 0, \text{ a positive}$$

Hence the distribution is that of the sum of a number of such variables, when the number is distributed in a Poisson Series of parameter m. The variate is, therefore, never negative, but is zero with finite frequency

$$e^{-m}.$$

Over the range of positive values, the distribution is continuous, and can be expressed as

$$\sqrt{\frac{m}{xa}}\, e^{-m-(x/a)} I_1\left(2\sqrt{\frac{xm}{a}}\right) dx$$

where I_1 is a Bessel function, specified by

$$I_1(u) = \tfrac{1}{2}u + \frac{1}{2^2.4}u^3 + \frac{1}{2^2.4^2.6}u^5 + \cdots$$

The distribution was first recognized (Bennett 1952, Fisher 1954) as characteristic of that of the length of germ plasm still heterogenic at an advanced stage of inbreeding, but its intrinsic incorporation of a finite condensation at

zero makes it appropriate to a number of natural phenomena, a good illustration being the rainfall of an arid region. A continuous model for rainfall is unsatisfactory for such regions. For many localities it is preferable to use a model ascribing the total rainfall for a given period, for example, a month or a year, to a number of showers, the number being a random sample from a Poisson series with parameter m, the rainfall of the showers having positive values only, representable by the Eulerian distribution

$$\frac{1}{p!}x^p e^{-x}\,dx$$

where p can be small, as in the previous example where it is actually 0, but must be > -1.

The advantage of this type of distribution, relevant for the purposes of the water engineer, is that there is a finite probability, namely e^{-m}, of no rain, whereas a continuous distribution would make this probability zero, contrary to experience.

For the asymptotic expansion, when m is sufficiently large, we may take the exact values for the mean (m) and variance ($2m$) and for the measures of non-normality

$$c = \frac{3}{\sqrt{2m}} \qquad f = \frac{90}{m^2}$$

$$d = \frac{6}{m} \qquad g = \frac{315\sqrt{2}}{m^2\sqrt{m}}$$

$$e = \frac{15\sqrt{2}}{m\sqrt{m}} \qquad h = \frac{2520}{m^3}.$$

These yield the six adjustments, to the normal deviate x

$$\text{I}\ \frac{x^2-1}{\sqrt{8m}} \qquad \text{IV}\ -\frac{4x^3-x}{384m^2}$$

$$\text{II}\ \frac{-x}{8m} \qquad \text{V}\ \frac{3x^4+2x^2-11}{480m^2\sqrt{2m}}$$

$$\text{III}\ \frac{x^2-1}{24m\sqrt{2m}} \qquad \text{VI}\ -\frac{96x^5+164x^3-767x}{46080m^3}.$$

The coefficients of Table IV give a rather comprehensive tabulation of the distribution, when m is sufficiently large for the accuracy required, and the levels of significance of interest. These large values of m would be troublesome to use in direct evaluation. For sufficiently small values of m, however, the probability that the variate x exceeds any limit X may be evaluated as

$$P = e^{-m}e^{-X}\sum_{i=0}^{\infty}\frac{X^i}{i!}\sum_{j>i}\frac{m^j}{j!}$$

which may be recognized as the probability that a random Poisson variate with parameter m shall exceed a random Poisson variate with parameter X.

TABLE IV—*Coefficients of powers of* $(2m)^{-1/2}$

p	x	$\frac{x^2-1}{2}$	$-\frac{x}{4}$	$\frac{x^2-1}{12}$	$-\frac{4x^3-x}{96}$	$\frac{3x^4+2x^2-11}{120}$	$-\frac{96x^5+164x^3-767x}{5760}$
.5	0	−0.50000	0	−0.08333	0	−0.09167	0
.25	0.67449	−0.27253	−0.16862	−0.04542	−0.00576	−0.07891	−0.09486
.10	1.28155	0.32119	−0.32039	0.05353	−0.07435	0.00314	0.05311
.05	1.64485	0.85277	−0.41121	0.14213	−0.16829	0.13642	−0.10835
.025	1.95996	1.42073	−0.48999	0.23679	−0.29330	0.34128	0.43543
.01	2.32635	2.20595	−0.58159	0.36766	−0.50035	0.73075	1.18428
.005	2.57583	2.81745	−0.64396	0.46957	−0.68527	1.11946	2.03348
.0025	2.80703	3.43972	−0.70176	0.57329	−0.89234	1.59180	3.16056
.001	3.09023	4.27477	−0.77256	0.71246	−1.19741	2.34733	−5.12555
.0005	3.29053	4.91378	−0.82263	0.81896	−1.45024	3.01970	7.00573

References

J. H. Bennett (1952). Junctions in inbreeding. *Genetica 26:* 392–406.

E. A. Cornish and R. A. Fisher (1937). Moments and cumulants in the specification of distributions. *Rev. de l'Inst. int. de stat. 5:* 307–22.

R. A. Fisher (1926). Expansion of "Student's" integral in powers of n^{-1}. *Metron 5:* 109–12.

R. A. Fisher (1941). The asymptotic approach to Behrens' integral, with further tables for the *d* test of significance. *Ann. eugen. 11:* 141–72.

R. A. Fisher (1954). A fuller theory of "junctions" in inbreeding. *Heredity 8:* 187–97.

R. A. Fisher (1957). The underworld of probability. *Sankhya 18:* 201–10.

R. A. Fisher (1959). "Statistical Methods and Scientific Inference." 2nd ed. Oliver and Boyd, Edin.

H. Goldberg and H. Levine (1946). Approximate formulas for the percentage points and normalization of *t* and χ^2. *Ann. math. statist. 17:* 216–25.

N. L. Johnson and B. L. Welch (1399). Applications of the non-central *t*-distribution. *Biometrika 31:* 362-89.

V. G. Panse (1940). A statistical study of quantitative inheritance. *Ann. eugen. 10:* 76–105.

P. V. Sukhatme (1938). On Fisher and Behrens' test of significance for the difference in means of two normal samples. *Sankhya 4:* 39–48.

An Approximation to the Negative Moments of the Positive Binomial Useful in Life Testing

W. MENDENHALL[1] AND E. H. LEHMAN, JR.

North Carolina State College

Raleigh, North Carolina

The purpose of this paper is to obtain the mean and variance of the maximum likelihood estimator of the scale parameter of a Weibull distribution where the sample is censored at a fixed time. It will be shown that these moments are functions of the negative moments of the positive binomial distribution. A simple approximation is obtained for the negative moments of the positive binomial, thus giving an approximate expression for the mean and variance of the estimator.

1. INTRODUCTION

A family of distributions useful in fitting failure populations is the function

$$f(t) = \left(\frac{m}{\alpha}\right) t^{m-1} e^{-t^m/\alpha} \quad \text{where} \quad t \geq 0, \alpha, m > 0. \tag{1.1}$$

It is known by engineers as the Weibull function as a result of its application by W. Weibull to several problems in engineering (10). The parameter m is called the shape parameter, the choice of which provides a wide range of distributional forms including the negative exponential ($m = 1$). The properties of the distribution as well as several methods of estimating the parameters, m and α, are discussed by Kao (7). In the case where the shape parameter is assumed known and we desire to estimate α, the maximum likelihood estimation of α reduces to the problem of estimating the parameters of the negative exponential distribution due to the invariance property of maximum likelihood estimators. Problems of estimation and tests concerning the parameter of a negative exponential distribution have been investigated by Epstein and Sobel (2, 3) and others. These results are applicable to the Weibull distribution when m is assumed known. An extensive list of references is given in (8).

Data obtained from a life test have the property that the observations are received in order of magnitude. The first item failing will be the smallest, the second will be the next to the smallest, and so on. We can take advantage of this property by concluding the test after either a fixed number of items fail or after a fixed length of time has elapsed. Such sampling procedures are called censored sampling. By increasing sample size, the length of time for a fixed number of units to fail will be reduced. One is therefore able to buy time at the expense of increased sample size. Censoring after a fixed number of items

[1]Now at Department of Mathematics, Bucknell University.

has failed has the advantage of providing a more or less uniform amount of information in repeated sampling with the disadvantage that the length of testing time varies from test to test. Censoring after a fixed length of testing time provides a constant length of test time in repeated tests with the amount of information varying from test to test. One advantage of the latter censoring procedure is that it simplifies the problem of test scheduling in a production process where information from periodic production lots must be obtained at regular time intervals.

In this paper we are interested in investigating the properties of the maximum likelihood estimator of the scale parameter, α, where it is assumed that m is known and that the sample is censored at a fixed time T. It can be shown that the maximum likelihood estimator of α is

$$\hat{\alpha} = \frac{\sum_{i=1}^{r} t_i^m + (n-r)T^m}{r} \tag{1.2}$$

where

n = number of items placed on test
T = length of test time
r = number failing before time T
$n - r$ = number surviving to time T
t_i = length of life of the ith ordered failure (where t_1 is the smallest, etc.)

In an experimental situation, the experimenter would like to know how large the sample size should be in order that the estimates fall within a given distance of the true value α. One method of expressing the "distance" numerically would be to use the mean square error, $E\{(\hat{\alpha} - \alpha)^2\}$, where E represents the operation of taking the average or expected value of the square of the difference $(\hat{\alpha} - \alpha)$. The problem may then be restated, "How large should the sample size be in order that $E\{(\hat{\alpha} - \alpha)^2\} = K$, where K is some desired value?" Or, we may ask the converse, "For a given sample size and censoring time, what will be the size of the mean square error?"

It will be noted that the mean square error is a function of the mean and variance of $\hat{\alpha}$:

$$E\{(\hat{\alpha} - \alpha)^2\} = V(\hat{\alpha}) + \{E(\hat{\alpha}) - \alpha\}^2 \tag{1.3}$$

where

$V(\hat{\alpha})$ = variance of $\hat{\alpha}$ and
$E(\hat{\alpha})$ = average or expected value of $\hat{\alpha}$.

We will derive an expression for the mean and variance of $\hat{\alpha}$, showing that both are functions of the negative moments of the positive binomial, i.e., functions of $E\{1/r^k\}$ where r is distributed as the truncated binomial,

$$P(r) = \frac{\binom{n}{r} p^r q^{n-r}}{\sum_{r=1}^{n} \binom{n}{r} p^r q^{n-r}}, \qquad r = 1, 2, \cdots, n. \tag{1.4}$$

A simple approximation to these negative moments will be derived and compared with some tabulated exact values for $k = 1, 2$.

2. The Mean and Variance of $\hat{\alpha}$

Let t equal the length of time to failure for an experimental unit and let the probability density function for t (failure distribution) be

$$f(t) = \left(\frac{m}{\alpha}\right) t^{m-1} e^{-t^m/\alpha}, \qquad t \geq 0 \tag{2.1}$$

A random sample of n units is tested until a predetermined time T at which time the test is concluded. Times to failure for r observations are observed where it will be noted that r is random. Thus we observe $t_1, t_2, \cdots t_r$ where t_j is the jth ordered observation. A total of r units failed and $(n - r)$ units survived.

For the sake of convenience and without loss of generality, we can measure time in units of size T. Then

$$x = \frac{t}{T},$$

and

$$f(x) = \frac{m}{\beta} x^{m-1} e^{-x^m/\beta}, \qquad \beta, m > 0, \tag{2.2}$$

where

$$x \geq 0 \quad \text{and} \quad \beta = \frac{\alpha}{T^m}$$

It can be shown that the maximum likelihood estimator of β is

$$\hat{\beta} = \left(\frac{1}{r}\right)\left[\sum_{i=1}^{r} x_i^m + (n - r)\right]. \tag{2.3}$$

It would follow that $\hat{\alpha} = T^m \hat{\beta}$.

In finding the mean and variance of $\hat{\beta}$, and hence $\hat{\alpha}$, we will restrict our results to the set of estimates for which $r > 0$. When $r = 0$, the maximum likelihood estimate is infinite. In order to obtain more information on α, either n or T (or both) should be increased.

The expected value of $\hat{\alpha}$ for the case $m = 1$ is given by Bartholomew (1). It can be shown that the expected value of $\hat{\beta}$ remains the same, regardless of the value of m and that

$$E(\hat{\beta}) = \beta - \frac{q}{p} + nE\left(\frac{1}{r}\right) - 1 \tag{2.4}$$

where

$$q = e^{-1/\beta}, \qquad p = 1 - q \quad \text{and} \quad r = 1, 2, \cdots, n.$$

The variance of $\hat{\beta}$ is equal to

$$V(\hat{\beta}) = V\left(\frac{1}{r}\sum_1^r x_i^m\right) + n^2 V\left(\frac{1}{r}\right) + 2n \operatorname{Cov.}\left(\frac{1}{r}\sum_1^r x_i^m, \frac{1}{r}\right) \tag{2.5}$$

Denote the expectation or variance of a quantity, given r, as E_r or V_r, respectively, and let $P(j)$ be the probability that r equals j. Then considering the first term of (2.5),

$$V_r\left(\frac{1}{r}\sum_1^r x_i^m\right) = \frac{1}{r} V(x^m), \qquad x \leq 1, \tag{2.6}$$

and

$$\begin{aligned} V\left(\frac{1}{r}\sum_1^r x_i^m\right) &= P(1)V(x^m) + P(2)\frac{V(x^m)}{2} \\ &\quad + \cdots + P(n)\frac{V(x^m)}{n} \\ &= V(x^m)\sum_1^n \frac{1}{r} P(r) \\ &= E\left(\frac{1}{r}\right)V(x^m) \end{aligned} \tag{2.7}$$

It can be shown that the variance of x^m, given $x \leq 1$, is equal to

$$V(x^m) = \beta^2 - q/p^2 \tag{2.8}$$

from which it follows that

$$V\left(\frac{1}{r}\sum_1^r x_i^m\right) = (\beta^2 - q/p^2)E(1/r) \tag{2.9}$$

The third term of (2.5) involves the covariance of $1/r \sum_1^r x^m$ and $1/r$.

$$\operatorname{Cov}\left(\frac{1}{r}\sum_1^r x_i^m, \frac{1}{r}\right) = E\left(\frac{1}{r^2}\sum_1^r x_i^m\right) - E\left(\frac{1}{r}\sum_1^r x_i^m\right)E\left(\frac{1}{r}\right) \tag{2.10}$$

Note that

$$E_r\left(\frac{1}{r}\sum_1^r x_i^m\right) = E(x^m) \tag{2.11}$$

and hence that

$$E\left(\frac{1}{r}\sum_1^r x_i^m\right) = \sum_{r=1}^n P(r)E_r\left(\frac{1}{r}\sum_1^r x_i^m\right) = E(x^m) \tag{2.12}$$

Also,

$$\begin{aligned} E\left(\frac{1}{r^2}\sum_1^r x_i^m\right) &= P(1)E_1(x^m) + \frac{P(2)}{2}E_2\left(\frac{1}{2}\sum_1^2 x_i^m\right) \\ &\quad + \cdots + \frac{P(n)}{n}E_n\left(\frac{1}{n}\sum_1^n x_i^m\right) \end{aligned} \tag{2.13}$$

Utilizing (2.11),

$$E\left(\frac{1}{r^2}\sum_1^r x_i^m\right) = E(x^m)E\left(\frac{1}{r}\right) \tag{2.13}$$

Substituting (2.12) and (2.14) into (2.10), we find that the

$$\text{Cov}\left(\frac{1}{r}\sum_{1}^{r} x_i^m, \frac{1}{r}\right) = 0$$

and

$$\begin{aligned} V(\hat{\beta}) &= V\left(\frac{1}{r}\sum_{1}^{r} x^m\right) + n^2 V\left(\frac{1}{r}\right) \\ &= E\left(\frac{1}{r}\right)V(x^m) + n^2 V\left(\frac{1}{r}\right) \qquad (2.15) \\ &= n^2 V\left(\frac{1}{r}\right) + (\beta^2 - q/p^2)E\left(\frac{1}{r}\right) \end{aligned}$$

Summarizing; the mean, bias, and variance of $\hat{\beta}$ are

$$E(\hat{\beta}) = \beta - q/p + nE\left(\frac{1}{r}\right) - 1 \qquad (2.16)$$

Bias of $\hat{\beta} = E(\hat{\beta}) - \beta$

$$= nE\left(\frac{1}{r}\right) - q/p - 1 \qquad (2.17)$$

$$V(\hat{\beta}) = n^2 V\left(\frac{1}{r}\right) + (\beta^2 - q/p^2)E\left(\frac{1}{r}\right) \qquad (2.18)$$

It follows that

$$E(\hat{\alpha}) = T^m E(\hat{\beta}), \qquad (2.19$$

$$V(\hat{\alpha}) = T^{2m} V(\hat{\beta}), \qquad (2.20$$

and Bias $\hat{\alpha} = T^m$ (Bias of $\hat{\beta}$) (2.21)

The mean square error of $\hat{\beta}$ is given by the equation

$$E(\hat{\beta} - \beta)^2 = V(\hat{\beta}) + (\text{Bias } \hat{\beta})^2 \qquad (2.22)$$

and it would follow from (2.20) and (2.21) that

$$E(\hat{\alpha} - \alpha)^2 = T^{2m} E(\hat{\beta} - \beta)^2. \qquad (2.23)$$

It is apparent that the mean square errors of $\hat{\beta}$ and $\hat{\alpha}$ are functions of the moments of $1/r$ since

$$V\left(\frac{1}{r}\right) = E\left(\frac{1}{r^2}\right) - \left[E\left(\frac{1}{r}\right)\right]^2 \qquad (2.24)$$

and can be computed given n, T, an approximate value of α, and the first two moments of $1/r$. Desk computation of $E(1/r)$ and $E(1/r^2)$ is exceedingly laborious. To overcome this difficulty, we present an abbreviated set of tables in Section 3 and a simple approximation formula in Section 4.

3. Negative Moments of the Positive Binomial Distribution

Examples of the use of the negative moments of the positive binomial distribution are given in papers by Stephan (9) and Grab and Savage (5). The latter paper discusses in particular the situation arising in sample survey theory where it is desired to find the variance of a sample mean,

$$\bar{y} = \frac{\sum_{1}^{r} y_i}{r} \quad \text{where the} \quad y_i\,, \qquad i = 1, 2, \cdots r,$$

are identically distributed with mean equal to μ, variance equal to σ^2 and r is a random variable distributed as a positive binomial variate. The variance of $\bar{y}$ is given by

$$V(\bar{y}) = \sigma^2 E\left(\frac{1}{r}\right), \quad \text{thus requiring}$$

the first negative moment of r. Reference is made to extensive discussions of this problem in (4) and (6). A table of $E(1/r)$ is given in (5) for

$$p = .01,\ .05(.05).95,\ .99 \quad \text{and} \quad n = 2(1)20.$$

Table 1.

Negative Moments of Positive Binomial Distribution $E(1/r^k)$

	$k = 1$					
n / 100 p	5	10	15	20	30	40
5	.94920	.88767	.82859	.77222	.66830	.57680
10	.89696	.77768	.67069	.57691	.42811	.32381
15	.84357	.67312	.53411	.42557	.28174	.20130
20	.78940	.57682	.42284	.31750	.19911	.14190
25	.73489	.49095	.33667	.24403	.15157	.10937
30	.68055	.41674	.27228	.19472	.12215	.08909
35	.62697	.35440	.22502	.16104	.10239	.07523
40	.57474	.30327	.19034	.13717	.08821	.06512
45	.52451	.26204	.16453	.11954	.07752	.05742
50	.47688	.22911	.14486	.10599	.06915	.05136
55	.43241	.20285	.12945	.09525	.06243	.04645
60	.39156	.18177	.11707	.08650	.05690	.04241
65	.35465	.16467	.10689	.07925	.05227	.03901
70	.32183	.15057	.09837	.07312	.04835	.03612
75	.29311	.13876	.09112	.06788	.04497	.03362
80	.26829	.12872	.08488	.06335	.04203	.03145
85	.24704	.12006	.07945	.05938	.03946	.02955
90	.22891	.11252	.07467	.05588	.03718	.02786
95	.21340	.10589	.07044	.05278	.03515	.02635

TABLE 1—*Continued*

n / $100p$	$k = 2$					
	5	10	15	20	30	40
5	.92438	.83489	.75122	.67352	.53616	.42214
10	.84785	.68013	.53810	.42095	.25209	.14993
15	.77099	.54003	.36857	.24789	.11260	.05498
20	.69446	.41791	.24342	.14155	.05290	.02453
25	.61905	.31567	.15720	.08136	.02811	.01353
30	.54559	.23352	.10116	.04881	.01708	.00863
35	.47499	.17014	.06622	.03130	.01152	.00601
40	.40815	.12303	.04488	.02155	.00834	.00444
45	.34594	.08915	.03184	.01577	.00634	.00342
50	.28915	.06540	.02365	.01209	.00498	.00272
55	.23840	.04902	.01831	.00959	.00403	.00221
60	.19410	.03775	.01464	.00782	.00332	.00183
65	.15642	.02991	.01201	.00650	.00279	.00154
70	.12526	.02434	.01005	.00549	.00238	.00132
75	.10023	.02026	.00854	.00470	.00205	.00114
80	.08069	.01717	.00736	.00407	.00178	.00099
85	.06582	.01477	.00640	.00356	.00157	.00088
90	.05470	.01285	.00563	.00314	.00139	.00078
95	.04637	.01129	.00498	.00279	.00124	.00070

and for

$$p = .01, .05(.05).50, .99 \quad \text{and} \quad n = 21(1)30.$$

In this paper we present tables of $E(1/r^k)$ for

$$p = .05\ (.05)\ .95, n = 5, 10, 15, 20, 30, 40, \quad \text{and}$$
$$k = 1, 2, 3, 4.$$

Table 1 presents the negative moments for $k = 1, 2$; Table 2 for $k = 3, 4$.

The tabulated results are obtained from the relation

$$E\left(\frac{1}{r^k}\right) = \frac{1}{1 - q^n} \sum_{r=1}^{n} \frac{1}{r^k} \binom{n}{r} p^r q^{n-r}$$

and are rounded to five decimal places.

4. An Approximation to the Negative Moments of the Positive Binomial Distribution

Several approximations have been given for negative moments of the positive binomial distribution. Stephan (9) achieves an approximation by expanding $E(1/r)$ in a series of inverse factorials which converge to the true value of $E(1/r)$ as more and more terms are added. An extension of this result is utilized to obtain the higher negative moments. Stephan's procedure permits the calcu-

lation of $E(1/r^k)$ to any degree of accuracy and enables one to calculate upper and lower bounds for the exact value. Its disadvantage is that the formula for $E(1/r^k)$ lacks computational simplicity.

Numerous references to approximations for $E(1/r)$ are given in (5). Specifically, Grab and Savage mention the approximation

$$E\left(\frac{1}{r}\right) \approx \frac{1}{np - q} \tag{4.1}$$

and state that one can expect two place accuracy for values of $np > 10$. We will derive an approximation which is nearly as simple as (4.1), which can be used to approximate the negative moments for $k \geq 1$, and which appears to be more accurate than (4.1) for values of $np > 5$.

As previously mentioned, Stephan derives his approximation by expanding $1/r$ into a series and then finding the expected value of the first t terms. In other words, he obtains the expected value of an approximation to $1/r$. Rather than approximate $1/r$, we will approximate the probability distribution function of r and then find the exact value of $E(1/r)$ for the approximating distribution. In consideration of the moments desired, the Beta function is an obvious choice.

The Beta function is given by the equation

$$f(z) = \frac{z^{a-1}(1 - z)^{b-1}}{B(a, b)} \tag{4.2}$$

where

$$B(a, b) = \frac{(a - 1)!\,(b - 1)!}{(a + b - 1)!}\,; a, b > 0, \quad \text{and} \quad 0 \leq z \leq 1.$$

Let $y = r/n$ where r is the positive binomial variate. Then the probability distribution of y will be a histogram ranging from $1/2n$ to $(2n + 1)/2n$. We will approximate this function with the Beta function, $f(z)$, although obviously an error will be introduced because of the discrepancy between the ranges of y and z. This is negligible for large values of np.

The parameters of the approximating Beta function were obtained by equating the first two moments of the two distributions. The two equations are:

$$\frac{na}{(a + b)} = np/(1 - q^n) \tag{4.3}$$

and

$$n^2 \frac{a(a + 1)}{(a + b)(a + b + 1)} = (npq + n^2p^2)/(1 - q^n)$$

where the moments of $P(r)$ are obtained from (9). In order to simplify the solution for a and b, let us assume that np is large and hence that $(1 - q^n)$ is approximately equal to one. The values of a and b are then the simultaneous solution of

$$\frac{na}{a+b} = np$$

and (4.3)

$$\frac{n^2a(a+1)}{(a+b)(a+b+1)} = npq + n^2p^2$$

Solving,

$$a = (n-1)p$$
$$b = (n-1)q \tag{4.4}$$

The approximating distribution is then

$$f(y) = \frac{y^{a-1}}{B(a,b)}(1-y)^{b-1} \tag{4.5}$$
$$= \frac{y^{(n-1)p-1}(1-y)^{(n-1)q-1}}{B(a,b)} \quad \text{where} \quad 0 \le y \le 1.$$

The negative moments of y are easily obtained from (4.5) yielding

$$E\left(\frac{1}{y^k}\right) = \frac{B(a-k,b)}{B(a,b)} \tag{4.6}$$

$$= \frac{(a-k-1)!(b-1)!}{(a+b-k-1)!} \Big/ \frac{(a-1)!(b-1)!}{(a+b-1)!} \tag{4.7}$$

$$= \frac{(a+b-1)(a+b-2)\cdots(a+b-k)}{(a-1)(a-2)\cdots(a-k)}$$

Specifically, from (4.7),

$$E\left(\frac{1}{y}\right) = \frac{(a+k-1)}{a-1} = \frac{n-2}{a-1} \tag{4.8}$$

and

$$E\left(\frac{1}{y^2}\right) = \frac{(n-2)(n-3)}{(a-1)(a-2)} \tag{4.9}$$

As noted in (4.1), in order for the moments (4.6) to exist, it is necessary that

$$a = (n-1)p > k. \tag{4.10}$$

The approximations to the moments of $1/r$ are easily obtained from the moments of $1/y$. Since $1/r = 1/(ny)$

$$E\left(\frac{1}{r^k}\right) \approx \frac{1}{n^k} E\left(\frac{1}{y^k}\right) \tag{4.11}$$

In particular,

$$E\left(\frac{1}{r}\right) \approx \frac{n-2}{n(a-1)} \tag{4.12}$$

and

$$E\left(\frac{1}{r^2}\right) \approx \frac{(n-2)(n-3)}{n^2(a-1)(a-2)} \tag{4.13}$$

Utilizing (4.12) and (4.13), it can be shown that

$$V\left(\frac{1}{r}\right) \approx \frac{(n-2)(n-a-1)}{n^2(a-1)^2(a-2)} \tag{4.14}$$

In order to distinguish these approximations from others, we will call them the Beta approximations.

The accuracy of these approximations will be discussed in Section 5. It might be well, however, to comment upon the errors introduced by our simplifying assumptions. The first error was caused by the error in approximating the range of the histogram of $P(r)$ as given by equation (1.4). The second was caused by the assumption that q^n is small and hence that the equations for the solution of a and b are approximately as given in (4.3). Both assumptions were made to obtain simplicity in the approximation and both errors diminish as np increases. We will see in Section 5 that the approximation is quite good for moderate sized values of np in spite of the assumptions.

5. A Comparison of the Approximation With Tabulated Values of $E(1/r)$ and $E(1/r^2)$

A comparison of the approximation to $E(1/r)$ with the tabulated values for various combinations of n and p is presented in Table 3. In addition, the approximation (4.1) is included. Proceeding from top to bottom, the three values in each cell are the tabulated value, the Beta approximation (4.12), and the approximation (4.1), respectively. All values are rounded in the sixth decimal place. The tabulated value for $n = 100$, $p = .90$ is omitted due to computing difficulties. Also, no approximation is given for the cells where $a \leq k$.

Examining Table 3, we note that the Beta approximation is correct to two decimal places for $np \geq 5$ and in several instances is correct to six decimal places. The approximation (4.1) also appears to be quite good although the Beta approximation would, in general, appear to be more accurate and hence more suitable for use in computing the mean square error of the estimator $\hat{a}$ mentioned in Section 2. As n and p increase, $(a-1) \to (n-2)$, $q \to 0$, and as a result, both the Beta approximation and the approximation given by equation (4.1) approach $1/n$.

Stephan (8) gives two examples of the use of his approximation. In the first example, $n = 100$, $p = .1$ and $E(1/r) = .111527$, correct to six decimal places. Utilizing two terms of the series, he obtains $E(1/r) \approx .108675$; for three terms, $E(1/r) \approx .110548$. The Beta approximation is .110112. In the second example, $n = 1000$, $p = .3$ and $E(1/r)$, correct to nine places, is .00334116. Two terms of the series yield $E(1/r) \approx .00334108$ and for three terms, $E(1/r) \approx .00334116$. The Beta approximation is .00334114. Thus, in these two examples, it requires three terms of Stephan's series to equal the accuracy of the Beta approximation.

The tabulated value of $E(1/r^2)$ and the Beta approximation are given in

TABLE 2

Negative Moments of Positive Binomial Distribution $E(1/r^k)$

	$k = 3$					
n / 100 p	5	10	15	20	30	40
5	.91216	.80959	.71513	.62878	.47972	.36014
10	.82408	.63555	.48109	.35810	.19045	.09761
15	.73651	.48239	.30380	.18559	.06578	.02326
20	.65028	.35318	.18042	.08892	.02152	.00604
25	.56636	.24903	.10141	.04046	.00745	.00206
30	.48574	.16905	.05459	.01818	.00303	.00094
35	.40946	.11061	.02867	.00848	.00149	.00052
40	.33854	.07000	.01508	.00430	.00086	.00032
45	.27391	.04314	.00821	.00244	.00055	.00021
50	.21634	.02619	.00476	.00153	.00038	.00015
55	.16639	.01593	.00299	.00104	.00027	.00011
60	.12435	.00991	.00202	.00074	.00020	.00008
65	.09015	.00643	.00144	.00055	.00015	.00006
70	.06345	.00440	.00108	.00042	.00012	.00005
75	.04352	.00318	.00083	.00033	.00009	.00004
80	.02942	.00240	.00065	.00027	.00008	.00003
85	.01999	.00187	.00052	.00022	.00006	.00003
90	.01401	.00149	.00043	.00018	.00005	.00002
95	.01034	.00121	.00035	.00015	.00004	

	$k = 4$					
n / 100 p	5	10	15	20	30	40
5	.90611	.79729	.69792	.60784	.45436	.33346
10	.81245	.61457	.45536	.33099	.16640	.07942
15	.71985	.45627	.27647	.16130	.05059	.01494
20	.62925	.32509	.15588	.07080	.01341	.00255
25	.54167	.22155	.08167	.02833	.00332	.00048
30	.45820	.14399	.03987	.01053	.00085	.00013
35	.37993	.08902	.01825	.00374	.00026	.00005
40	.30790	.05225	.00793	.00134	.00011	.00002
45	.24305	.02909	.00334	.00052	.00005	.00001
50	.18612	.01539	.00142	.00024	.00003	.00001
55	.13758	.00779	.00065	.00012	.00002	.0000
60	.09760	.00384	.00033	.00008	.00001	.00000
65	.06599	.00190	.00019	.00005	.00001	.00000
70	.04221	.00099	.00012	.00003	.00001	.00000
75	.02535	.00057	.00008	.00002	.00000	.00000
80	.01426	.00036	.00006	.00002	.00000	.00000
85	.00762	.00025	.00004	.00001	.00000	.00000
90	.00408	.00018	.00003	.00001	.00000	.00000
95	.00240	.00013	.00002	.00001	.00000	

TABLE 3. *A Comparison of the Tabulated Value of $E(1/r)$, the Beta Approximation, and the Approximation Given by Equation (4.1)*
Entries: Tabulated Value
Beta Approximation
Approximation (4.1)

p \ n	5	10	20	30	40	100
.10			.576912	.428114	.323807	.111527
			1.00000	.491228	.327586	.110112
			.909091	.476190	.322581	.109890
.30	.680554	.416735	.194724	.122147	.089094	.034159
	3.00000	.470588	.191489	.121212	.088785	.034146
	1.25000	.434783	.188679	.120482	.088496	.034130
.50	.476882	.229109	.105990	.069152	.051356	.020206
	.600000	.228571	.105882	.069136	.051351	.020206
	.500000	.222222	.105263	.068966	.051282	.020202
.70	.321834	.150568	.073123	.048345	.036116	.014348
	.333333	.150943	.073171	.048359	.036122	.014348
	.312500	.149254	.072993	.048309	.036101	.014347
.90	.228913	.112523	.055885	.037180	.027857	
	.230769	.112676	.055901	.037185	.027859	.011124
	.227273	.112360	.055866	.037175	.027855	.011123

TABLE 4. *A Comparison of the Tabulated Value of $E(1/r^2)$ and the Beta Approximation*
Entries: Tabulated Value
Beta Approximation

p \ n	5	10	20	30	40	100
.1				.252093	.149937	.014559
				.049123	.159483	.013520
.3		.233527	.048813	.017077	.008627	.001198
		.470588	.043991	.016282	.008467	.001196
.5		.065403	.012090	.004985	.002716	.000413
		.064000	.012000	.004978	.002714	.000413
.7	.125262	.024343	.005489	.002375	.001320	.000207
	.166667	.024572	.005504	.002378	.001321	.000207
.9	.054698	.012847	.003143	.001388	.000778	
	.057692	.012930	.003147	.001389	.000779	.000124

Table 4, the tabulated value followed by the approximation. The agreement appears to be quite good for moderate sized values of np.

6. An Example of the Determination of the Mean Square Error of $\hat{\alpha}$

We plan to conduct a life test in which time, T, and sample size, n, are dictated by the amount and availability of material and equipment. We desire a measure of the goodness of estimation which the money represented by n and T will buy. As mentioned in Section 1, the mean square error of the estimator provides this measure.

For example, let us assume that we wish to estimate the parameter α of the Weibull distribution for the estimation problem discussed in Sections 1 and 2. Further, let us assume that from previous experience, we know that the failure distribution has a form represented by the Weibull distribution with shape parameter $m = 2$ and that the mean life lies somewhere in the range of 80 to 120 hours. We might choose $T = 200$ hours and sample size $n = 40$.

In calculating the mean square error of $\hat{\alpha}$, we will use equations (2.17, 2.18, and 2.22) to obtain the mean square error of $\hat{\beta}$. Utilizing (2.23), we can determine the mean square error of $\hat{\alpha}$.

Looking at equations (2.17 and 2.18), we see that we need to know the values $E(1/r)$, $V(1/r)$, β, p and q. Since β (or α) is unknown and is the parameter we wish to estimate, the best that we can do is to substitute a guessed value for β.

The mean value of t for the Weibull distribution is

$$E(t) = \alpha^{1/m}\Gamma\left(\frac{1}{m} + 1\right) \tag{6.1}$$

or

$$\alpha = \left[\frac{E(t)}{\Gamma\left(\frac{1}{m} + 1\right)}\right]^m \tag{6.2}$$

Since we expect the mean life to fall in the range 80 to 120 hours, we might choose $E(t) = 100$ as a guessed value. Then, for our example, the guessed value of α would be

$$\alpha = \left[\frac{100}{\Gamma\left(\frac{1}{2} + 1\right)}\right]^2$$

Tables of $\Gamma(x)$ can be found in most standard handbooks of mathematical tables. We note $\Gamma(1.5) = .88623$. Then

$$\alpha \approx \left(\frac{100}{.88623}\right)^2 \approx 12{,}732 \tag{6.3}$$

From (2.2) and (2.4)

$$\beta = \frac{\alpha}{T^m} = \frac{12{,}732}{(200)^2} = .31830 \tag{6.4}$$

$$q = e^{-1/\beta} = e^{-1/.31830} = .04320 \tag{6.5}$$

and

$$p = 1 - q = .95680 \tag{6.6}$$

The values of $E(1/r)$ and $V(1/r)$ can be determined from the Beta approximations (4.12) and (4.14).

$$E\left(\frac{1}{r}\right) = \frac{n-2}{n(a-1)} \quad \text{where} \quad a = (n-1)p = (39)(.9568) = 37.315$$

$$E\left(\frac{1}{r}\right) = \frac{38}{40(36.315)} = .0262$$

$$V\left(\frac{1}{r}\right) = \frac{(n-2)(n-a-1)}{n^2(a-1)^2(a-2)}$$

$$= \frac{(38)(40-37.315-1)}{(40)^2(36.315)^2(35.315)} = .00000085927$$

Substituting the values into (2.17) and (2.18), Bias $\hat{\beta} = nE(1/r) - q/p - 1$

$$= 40(.0262) - \frac{.0432}{.9568} - 1$$

$$= .0013$$

$$V(\hat{\beta}) = n^2 V\left(\frac{1}{r}\right) + (\beta^2 - q/p^2)E\left(\frac{1}{r}\right)$$

$$= (40)^2(.0000009) + (.318)^2 - \frac{.0432}{(.9568)^2}(.0262)$$

$$= .00284$$

The mean square error of $\hat{\beta}$ and $\hat{\alpha}$ can be calculated from (2.22) and (2.23) respectively.

$$E\{(\hat{\beta} - \beta)^2\} = V(\hat{\beta}) + (\text{Bias } \hat{\beta})^2$$

$$= .00284 + (.0013)^2 = .002849$$

In this case, the square of the bias is so small compared with the variance that it can be neglected.

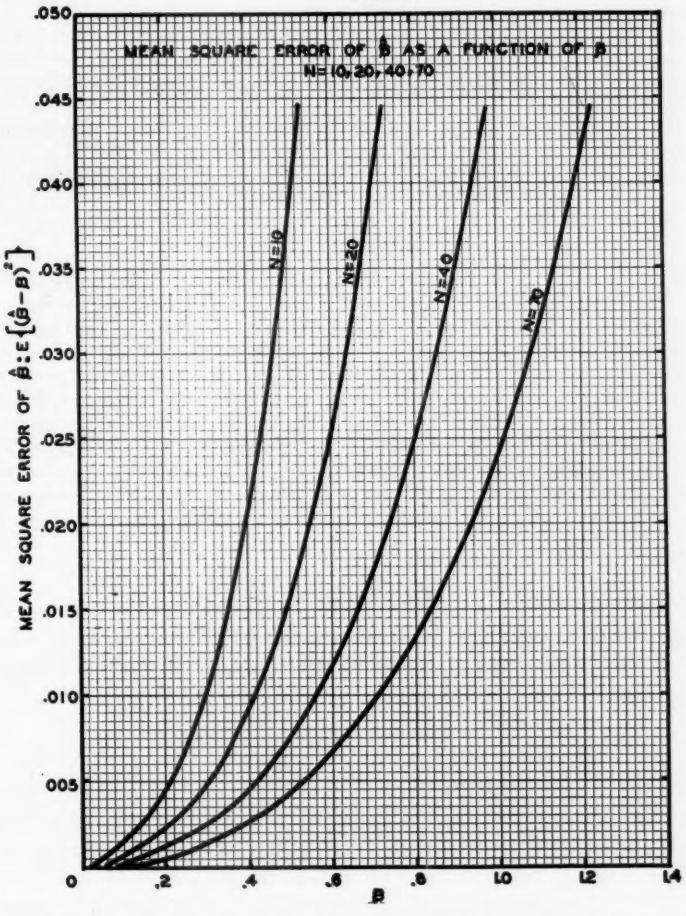
$$E\{(\hat{\alpha} - \alpha)^2\} = T^{2m}E\{(\hat{\beta} - \beta)^2\}$$

$$= (200)^4(.00285) = 4{,}574{,}000$$

When the bias equals zero, the mean square error equals the variance and consequently, for this case (bias approximately equal to zero), the standard deviation of the estimator, $\hat{\alpha}$, is 2139.

The mean square error as a function of β is presented in Figure 1 for $n = 10$,

FIGURE 1



20, 40, and 70. For the above example, we read from Figure 1 ($n = 40$) the mean square error of β is approximately .0030 for $\hat{\beta} = .318$.

The authors wish to thank Dr. A. H. E. Grandage for assistance in the preparation of Tables 1 and 2.

References

1. Bartholomew, D. J. 1957. A problem in life testing. J. Amer. Stat. Assn. *52*: 350–355.
2. Epstein, B. and Sobel, M. 1953. Life testing. J. Amer. Stat. Assn. *48*: 486–502.
3. Epstein, B. and Sobel, M. 1954. Some theorems relevant to life testing from an exponential distribution. Ann. Math. Stat. *25*:373–381.
4. Finkner, A. L., 1950. Further investigation on the theory and application of sampling for scarcity items. Institute of Statistics, University of North Carolina, Mimeo. Series 30.
5. Grab, E. L. and Savage, I. R. 1954. Tables of the expected value of $1/X$ for positive Bernoulli and Poisson variables. J. Amer. Stat. Assn. *49*: 169–177.

6. HORVITZ, D. G. 1953. Ratio method of estimation in sample surveys. Unpublished thesis, Iowa State College.
7. KAO, J. H. K. 1955. The Weibull distribution in life testing of electron tubes. Unpublished paper presented at annual meeting of American Statistical Association, New York.
8. MENDENHALL, W. 1958. A bibliography on life testing and related topics. Biometrika *45*: 521–543.
9. STEPHAN, F. F. 1945. The expected value and variance of the reciprocal and other negative powers of a positive Bernoullian variate. Ann. Math. Stat. *16*: 50–61.
10. WEIBULL, W. 1951. A statistical distribution function of wide applicability. J. Appl. Mech. *18*: 293–297.

Order Statistics from the Gamma Distribution

SHANTI S. GUPTA

Bell Telephone Laboratories, Incorporated

Allentown, Pennsylvania

This paper deals with order statistics from a gamma or χ^2 (Pearson Type III) distribution. Expressions are derived for the moments of an order statistic and for the covariance between two order statistics. A table of moments (about the origin as well as the mean) is presented. Equations are derived which are used to solve for the percentage points of the order statistics and a table of the percentage points is given. The problem of the best linear unbiased estimate of the scale parameter of the distribution which is based on order statistics is, briefly, discussed for the censored (type II) case. The modal value of an order statistic is derived and a table of these values is given for the cases of the smallest and the largest order statistics. Applications to life-tests, extreme values, reliability and maintenance are described and illustrated in some cases.

1. INTRODUCTION

Suppose that we have N independent random variables $x_1, x_2 \cdots, x_N$ each distributed according to the gamma probability density function (p.d.f.)

$$g_r(x) = e^{-x}\frac{x^{r-1}}{\Gamma(r)}, \qquad 0 \leq x < \infty \tag{1.1}$$

where r is assumed to be a positive integer. The distribution (1.1) arises, for instance, as the convolution of r independent and identical chance variables with the exponential density function e^{-x}. Let $G_r(x)$ be the cumulative distribution function of the x, then it is easy to see that $G_r(x)$ can be written as a partial sum of probabilities in a Poisson distribution as follows

$$G_r(x) = \sum_{j=r}^{\infty}\frac{e^{-x}x^j}{j!}. \tag{1.2}$$

Let $x_{(k)}$ denote the kth smallest observation when the x's are arranged in the increasing order, viz.

$$x_{(1)} \leqq x_{(2)} \leqq \cdots \leqq x_{(k)} \leqq \cdots \leqq x_{(N)}, \tag{1.3}$$

and let $y = x_{(k)}$, then we are interested in the moments and distribution of y.

2. MOMENTS OF THE kth ORDER STATISTIC

The ith moment about the origin, $\mu_i'(k, N)$ of the kth order statistic from a sample of size N from the gamma p.d.f. $g_r(x)$ is given by

$$\mu_i'(k, N) = \frac{N!}{(k-1)!\,(N-k)!} \cdot \int_0^\infty \left[1 - \sum_{j=0}^{r-1} \frac{e^{-y}(y)^j}{j!}\right]^{k-1} \left[\sum_{j=0}^{r-1} \frac{e^{-y}(y)^j}{j!}\right]^{N-k} \frac{e^{-y} y^{r+i-1}}{\Gamma(r)}\, dy \tag{2.1}$$

which reduces to (details omitted)

$$\mu_i'(k, N) = \frac{N!}{(k-1)!\,(N-k)!\,\Gamma(r)} \sum_{p=0}^{k-1} (-1)^p \binom{k-1}{p} \cdot \sum_{m=0}^{(r-1)(N-k+p)} a_m(r, N-k+p) \frac{\Gamma(r+i+m)}{(N-k+p+1)^{r+i+m}} \tag{2.2}$$

where $a_m(r, p)$ is the coefficient of t^m in the expansion of $\left(\sum_{j=0}^{r-1} \frac{t^j}{j!}\right)^p$. For $k = 1$, we have from (2.2)

$$\mu_i'(1, N) = \frac{N}{\Gamma(r)} \sum_{m=0}^{(r-1)(N-1)} a_m(r, N-1) \frac{\Gamma(r+i+m)}{N^{r+i+m}}. \tag{2.3}$$

The ith moment of the kth order statistic can be expressed in terms of the ith moments of 1st (smallest) order statistic as follows

$$\mu_i'(k, N) = \frac{N!}{(k-1)!\,(N-k)!} \sum_{p=0}^{k-1} (-1)^p \binom{k-1}{p} \frac{\mu_i'(1, N-k+p+1)}{(N-k+p+1)}. \tag{2.4}$$

Also, the coefficients $a_m(r, p)$ satisfy the following basic recursion relation.

$$a_m(r, p) = a_m(r, p-1) + a_{m-1}(r, p-1) + \frac{1}{2!} a_{m-2}(r, p-1) + \cdots + \frac{1}{(r-1)!} a_{m-r+1}(r, p-1) \tag{2.5}$$

Since for $p = 1$, these coefficients are $1, 1, 1/2!, \cdots, 1/(r-1)!$, we can compute them for any value of p by repeated applications of (2.5).

If we define

$$\nu_i'(k, N) = \int_0^\infty x^i G_r^{k-1}(x)[1 - G_r(x)]^{N-k} g_r(x)\, dx = \frac{(k-1)!\,(N-k)!\,\Gamma(r)}{N!}\, \mu_i'(k, N), \tag{2.6}$$

then $\nu_i'(k, N)$ satisfy the following recursion formula

$$\nu_i'(k, N) = \sum_{\alpha=0}^{N-k} (-1)^\alpha \binom{N-k}{\alpha} \nu_i'(k+\alpha, k+\alpha). \tag{2.7}$$

A slightly more generalized form of (2.7) is the following one.

$$\nu_i'(k, N) = \sum_{p=0}^{\beta} (-1)^p \binom{\beta}{p} \nu_i'(k+p, N-\beta+p) \tag{2.8}$$

where β is any positive integer $\leqq N - k$. Also, the relation (2.4) can be written as

$$\nu_i'(k, N) = \sum_{p=0}^{k-1} (-1)^p \binom{k-1}{p} \nu_i'(1, N-k+p+1). \tag{2.9}$$

From (2.8), we have, by putting $\beta = 1$,

$$\nu_i'(k, N-1) = \nu_i'(k, N) + \nu_i'(k+1, N), \qquad k \leqq N-1. \tag{2.10}$$

Table I gives the first four moments (about the origin as well as about the mean) of the kth order statistic for selected values of k, N, and r.

TABLE I

The entries in this table are the moments about the origin $\mu_i'(k,N)$ and the central moments $\mu_i(k,N)$ of the k^{th} order statistic of a sample of size N from a standardized gamma distribution with parameter r.

r = 1

N	k	$\mu_1'(k,N)$	$\mu_2'(k,N)$	$\mu_3'(k,N)$	$\mu_4'(k,N)$	$\mu_2(k,N)$	$\mu_3(k,N)$	$\mu_4(k,N)$
1	1	1.00000	2.00000	6.00000	24.0000	1.00000	2.00000	9.00000
2	1	.500000	.500000	.750000	1.50000	.250000	.250000	.562500
	2	1.50000	3.50000	11.2500	46.5000	1.25000	2.25000	11.0625
3	1	.333333	.222222	.222222	.296296	.111111	.0740741	.111111
	2	.833333	1.05556	1.80556	3.90741	.361111	.324074	.840278
	3	1.83333	4.72222	15.9722	67.7963	1.36111	2.32407	1.20069
4	1	.250000	.125000	.0937500	.0937500	.0625000	.0312500	.0351563
	2	.583333	.513889	.607639	.903935	.173611	.105325	.187934
	3	1.08333	1.59722	3.00347	6.91088	.423611	.355324	1.01085
	4	2.08333	5.76389	20.2951	88.0914	1.42361	2.35532	12.5525
5	1	.200000	.0800000	.0478000	.0384000	.0400000	.0160000	.0144000
	2	.450000	.305000	.276750	.315150	.102500	.0472500	.0645563
	3	.783333	.827222	1.10397	1.78711	.213611	.121324	.244001
	4	1.28333	2.11056	4.26951	10.3267	.463611	.371324	1.12692
	5	2.28333	6.67722	24.3015	107.533	1.46361	2.37132	12.9086
6	1	.166667	.0555556	.0277778	.0185185	.0277778	.00925926	.00694444
	2	.366667	.202222	.149111	.137807	.0677778	.0252593	.0280111
	3	.616667	.510556	.532028	.669835	.130278	.0565093	.0885840
	4	.950000	1.14389	1.67592	2.90439	.241389	.130583	.286547
	5	1.45000	2.59389	5.56675	14.0379	.491389	.380583	1.21113
	6	2.45000	7.49389	28.0484	126.232	1.49139	2.38058	13.1595
7	1	.142857	.0408163	.0174927	.00999583	.0204082	.00583090	.00374844
	2	.309524	.143991	.0894882	.0696546	.0481859	.0150902	.0140942
	3	.509524	.347800	.298168	.308189	.0881859	.0310902	.0400589
	4	.759524	.727563	.843540	1.15203	.150686	.0623401	.108885
	5	1.09286	1.45613	2.29927	4.21866	.261798	.136414	.319854
	6	1.59286	3.04899	6.87346	17.9656	.511796	.386416	1.27504
	7	2.59286	8.23471	31.5776	144.276	1.51180	2.38641	13.3458
8	1	.125000	.0312500	.0117188	.00585938	.0156250	.00390625	.00219727
	2	.267857	.107781	.0579104	.0389511	.0360332	.00973716	.00785897
	3	.434524	.252622	.184221	.161765	.0638109	.0189964	.0208089
	4	.634524	.506431	.488080	.552229	.103811	.0349965	.0505235
	5	.884524	.948694	1.19960	1.75183	.166311	.0662462	.124609
	6	1.21786	1.76060	2.96020	5.69876	.277422	.140321	.346592
	7	1.71786	3.47846	6.17788	22.0545	.527422	.390323	1.32522
	8	2.71786	8.91417	34.9204	161.736	1.52742	2.39031	13.4898
9	1	.111111	.0246914	.00823045	.00365798	.0123457	.00274348	.00137174
	2	.236111	.0837191	.0396251	.0234705	.0279707	.00664973	.00472642
	3	.378968	.191996	.121909	.0931328	.0483789	.0124806	.0118998
	4	.545635	.373874	.308548	.289030	.0761564	.0217399	.0269073
	5	.745634	.672127	.712122	.868728	.116158	.0377396	.0595857
	6	.995636	1.16995	1.58958	2.45831	.178657	.0689887	.138302
	7	1.32897	2.05592	3.64550	7.31898	.289774	.143057	.368531
	8	1.82897	3.88489	9.47285	26.2647	.539762	.393073	1.36564
	9	2.82897	9.54283	38.1013	178.670	1.53977	2.39306	13.6043
10	1	.100000	.0200000	.00600000	.00240000	.0100000	.00200000	.000900000
	2	.211111	.0669136	.0283045	.0149798	.0223457	.00474348	.00301248
	3	.336111	.150941	.0849075	.0574336	.0379707	.00864973	.00730466
	4	.478968	.287789	.208246	.176431	.0583790	.0144806	.0157026
	5	.645637	.503002	.459747	.482929	.0861552	.0237405	.0323758
	6	.845632	.841253	.964498	1.25453	.126159	.0397384	.0674573
	7	1.09564	1.38908	2.00631	3.26083	.188650	.0709957	.149908
	8	1.42896	2.34171	4.34802	9.05819	.299783	.145045	.386844
	9	1.92897	4.27069	10.7541	30.5663	.549752	.395096	1.39885
	10	2.92897	10.1286	41.1399	195.126	1.54977	2.39504	13.6977
11	1	.0909091	.0165289	.00450789	.00163923	.00826446	.00150263	.000614712
12	1	.0833333	.0138889	.00347222	.00115741	.00694444	.00115741	.000434028
13	1	.0769231	.0118343	.00273100	.000840307	.00591716	.000910332	.000315115
14	1	.0714286	.0102041	.00218659	.000624740	.00510204	.000728863	.000234277
15	1	.0666667	.00888889	.00177778	.000474074	.00444444	.000592593	.000177778

TABLE 1 (Cont.)

The entries in this table are the moments about the origin $\mu_1'(k,N)$ and the central moments $\mu_1(k,N)$ of the k^{th} order statistic of a sample of size N from a standardized gamma distribution with parameter r.

r = 2

N	k	$\mu_1'(k,N)$	$\mu_2'(k,N)$	$\mu_3'(k,N)$	$\mu_4'(k,N)$	$\mu_2(k,N)$	$\mu_3(k,N)$	$\mu_4(k,N)$
1	1.	2.00000	6.00000	24.00000	120.000	2.00000	4.00000	24.0000
2	1	1.25000	2.25000	5.25000	15.0000	.687500	.718750	2.51953
	2	2.75000	9.75000	42.7500	225.000	2.18750	3.90625	25.5820
3	1	.962963	1.30864	2.27160	4.77366	.381344	.276990	.725138
	2	1.82407	4.13272	11.2068	35.4527	.805470	.729937	2.97629
	3	3.21296	12.5586	58.5216	319.774	2.23551	3.80592	25.8286
4	1	.804688	.902344	1.28027	2.18262	.254822	.144065	.309607
	2	1.43779	2.52754	5.24560	12.5468	.460299	.287910	.908231
	3	2.21036	5.73790	17.1680	58.3586	.852209	.717797	3.10051
	4	3.54716	14.8322	72.3061	406.912	2.24985	3.73264	25.7841
5	1	.702080	.680832	.829747	1.20846	.187916	.0878850	.162927
	2	1.21512	1.78839	3.08238	6.07923	.311880	.151331	.400641
	3	1.77180	3.63625	8.49043	.222481	.496991	.286591	1.00093
	4	2.50273	7.13899	22.9530	82.4322	.875313	.704672	3.24801
	5	3.80827	16.7555	84.6444	488.032	2.25260	3.67779	25.6570
6	1	.629115	.543038	.586077	.752743	.147252	.0591656	.0975264
	2	1.06690	1.36980	2.04810	3.48706	.231516	.0926435	.214820
	3	1.51155	2.62557	5.15094	11.2636	.340804	.151984	.452531
	4	2.03205	4.64694	11.8299	33.2327	.517717	.283072	1.05480
	5	2.73808	8.38502	28.5146	107.032	.887952	.693181	3.29176
	6	4.02231	18.4296	95.8704	564.232	2.25065	3.63504	25.5079
7	1	.574020	.449720	.438746	.507695	.120221	.0425787	.0636846
	2	.959688	1.10295	1.47006	2.22304	.181949	.0623495	.130002
	3	1.33494	2.03693	3.49319	6.64713	.254851	.0935796	.246657
	4	1.74701	3.41044	7.36127	17.4188	.358379	.151010	.485760
	5	2.24582	5.57431	1.51814	45.0931	.530588	.279272	1.08836
	6	2.93498	9.50930	33.8478	131.807	.895201	.683445	3.31301
	7	4.20353	19.9164	106.207	636.303	2.24667	3.60052	25.3575
8	1	.530627	.382657	.342482	.362569	.101092	.0321495	.0442693
	2	.877767	.919162	1.11259	1.52357	.148686	.0447605	.0854290
	3	1.20545	1.65431	2.54247	4.32143	.201205	.0632010	.150933
	4	1.55077	2.67462	5.07771	10.5233	.269732	.0934092	.255349
	5	1.94326	4.14625	9.64484	24.3144	.370004	.149576	.508297
	6	2.42736	6.43115	18.5033	57.5604	.539062	.275637	1.11020
	7	3.10418	10.5354	38.9627	156.556	.899392	.675225	3.32150
	8	4.36058	21.2565	115.814	704.838	2.24182	3.57189	25.2129
9	1	.495368	.332304	.275891	.270309	.0869142	.0251688	.0322542
	2	.812698	.785479	.875207	1.10065	.125000	.0330747	.0595901
	3	1.10551	1.38705	1.94345	3.00379	.168900	.0454523	.0999522
	4	1.40533	2.18884	3.74053	6.95671	.213887	.0633354	.165764
	5	1.73257	3.28184	6.74917	14.9815	.280050	.0927456	.284092
	6	2.11181	4.83779	11.9614	31.7807	.378035	.148208	.523713
	7	2.58514	7.22783	21.7743	70.4602	.544901	.272124	1.12557
	8	3.25248	11.4804	43.8736	181.158	.901715	.668335	3.32221
	9	4.49909	22.4785	124.806	770.298	2.23667	3.54749	25.0770
10	1	.466022	.293204	.227768	.208389	.0760282	.0202667	.0243755
	2	.759490	.684202	.708999	.827589	.107377	.0262505	.0434858
	3	1.02553	1.19059	1.54004	2.19290	.138873	.0342232	.0700865
	4	1.29213	1.84547	2.88473	4.89587	.175879	.0456272	.110631
	5	1.57513	2.70388	5.02424	10.0480	.222850	.0632411	.170003
	6	1.89001	3.85980	8.47411	19.9151	.287669	.0916570	.290466
	7	2.25967	5.48977	14.2862	39.6911	.383643	.147327	.533372
	8	2.72462	7.97270	24.9835	83.6327	.549158	.268479	1.13750
	9	3.38445	12.3573	48.5961	205.540	.902774	.662781	3.31650
	10	4.62294	23.6031	133.274	833.049	2.23149	3.52634	24.9818
11	1	.441125	.262023	.191767	.165015	.0674317	.0166913	.0189677
12	1	.419673	.236612	.164071	.133561	.0604869	.0140022	.0151161
13	1	.400950	.215530	.142265	.110091	.0547700	.0119277	.0122881
14	1	.384431	.197776	.124759	.0921527	.0499889	.0102929	.0101578
15	1	.369720	.182629	.110470	.0781598	.0459362	.00898107	.00851788

TABLE I (Cont.)

The entries in this table are the moments about the origin $\mu_1'(k,N)$ and the central moments $\mu_1(k,N)$ of the k^{th} order statistic of a sample of size N from a standardized gamma distribution with parameter r.

r = 3

N	k	$\mu_1'(k,N)$	$\mu_2'(k,N)$	$\mu_3'(k,N)$	$\mu_4'(k,N)$	$\mu_2(k,N)$	$\mu_3(k,N)$	$\mu_4(k,N)$
1	1	3.00000	12.0000	60.0000	360.000	3.00000	6.00000	45.0000
2	1	2.06250	5.43750	17.3438	64.6875	1.18359	1.24658	6.09810
	2	3.93750	18.5625	102.656	655.313	3.05859	5.47998	44.1108
3	1	1.68313	3.55007	8.91632	25.8436	.717150	.526995	2.08043
	2	2.82125	9.21236	34.1986	142.375	1.25294	1.13857	6.33804
	3	4.49563	23.2376	136.885	911.781	3.02690	5.20198	42.7113
4	1	1.46603	2.66093	5.69475	13.9487	.511679	.293454	1.01016
	2	2.33441	6.21747	18.5810	61.5284	.768009	.481233	2.22680
	3	3.30808	12.2073	49.8162	223.222	1.26385	1.07172	6.29760
	4	4.89148	26.9143	165.908	1141.30	2.98780	5.02768	41.5241
5	1	1.32126	2.14332	4.07151	8.80054	.397589	.188978	.589690
	2	2.04512	4.73138	12.1877	34.5413	.548865	.266498	1.09395
	3	2.76834	8.44661	28.1710	102.009	.782896	.453280	2.25802
	4	3.66791	14.7144	64.2463	304.031	1.26080	1.02637	6.20068
	5	5.19737	29.9643	191.323	1350.62	2.95170	4.90563	40.5508
6	1	1.21592	1.80390	3.11787	6.10502	.325431	.133069	.385094
	2	1.84796	3.84043	8.83968	22.2782	.425491	.170160	.640423
	3	2.43945	6.51328	18.8838	59.0675	.562367	.251264	1.12271
	4	3.09723	10.3799	37.4582	144.951	.787086	.433571	2.25428
	5	3.95325	16.8816	77.6403	383.571	1.25340	.993299	6.09410
	6	5.44619	32.5809	214.060	1544.03	2.91989	4.81383	39.7439
7	1	1.13483	1.56364	2.49998	4.51253	.275816	.0995137	.271080
	2	1.70251	3.24544	6.82522	15.6600	.346889	.118640	.417531
	3	2.21156	5.32789	13.8758	38.8237	.436909	.160387	.661699
	4	2.74331	8.09380	25.5611	86.0591	.568069	.240599	1.13155
	5	3.36268	12.0945	46.3810	189.120	.786941	.418554	2.23773
	6	4.18948	18.7964	90.1441	461.351	1.24466	.968062	5.98994
	7	5.65564	34.8783	234.713	1724.47	2.89200	4.74128	39.0637
8	1	1.06987	1.38428	2.07151	3.48990	.239657	.0777047	.201307
	2	1.58951	2.81921	5.49932	11.6709	.292654	.0878013	.292766
	3	2.04152	4.52416	10.8029	27.6271	.356360	.111734	.432685
	4	2.49196	6.66745	18.9973	57.4848	.442659	.153527	.671864
	5	2.99166	9.58015	32.1250	114.633	.570124	.232691	1.13021
	6	3.58528	13.6392	54.9346	233.811	.784907	.406140	2.21852
	7	4.39088	20.5155	101.881	537.198	1.23567	.948368	5.89038
	8	5.83632	36.9301	253.689	1894.08	2.86746	4.68173	38.4822
9	1	1.01630	1.24502	1.75921	2.79193	.212152	.0626746	.155587
	2	1.49843	2.49837	4.56992	9.07370	.253074	.0678503	.216294
	3	1.90829	3.94212	8.75221	20.7611	.300530	.0824941	.303851
	4	2.30797	5.68825	14.9043	41.3590	.361527	.107268	.440276
	5	2.72868	7.89145	24.1135	77.6420	.445743	.147696	.678157
	6	3.20204	10.8231	38.5341	144.227	.570020	.227538	1.11944
	7	3.77690	15.0472	63.1349	278.604	.782219	.394453	2.20519
	8	4.56630	22.0778	112.951	611.082	1.22672	.933537	5.79088
	9	5.99508	38.7867	271.281	2054.46	2.84573	4.63123	37.9806
10	1	.971114	1.13359	1.52275	2.29286	.190532	.0518412	.124023
	2	1.42298	2.24783	3.88734	7.28351	.222970	.0541728	.166211
	3	1.80026	3.50055	7.30024	16.2344	.259613	.0636031	.224435
	4	2.16037	4.97244	12.1401	31.3234	.305224	.0790100	.310466
	5	2.52936	6.76195	19.0507	56.4124	.364310	.104326	.442685
	6	2.92800	9.02093	29.1764	98.8716	.447750	.141120	.687557
	7	3.38473	12.0245	44.7725	174.463	.568126	.226794	1.09111
	8	3.94497	16.3426	71.0045	323.235	.779836	.380449	2.21309
	9	4.72163	23.5116	123.437	683.044	1.21785	.923725	5.68290
	10	6.13657	40.4839	287.708	2206.84	2.82640	4.58707	37.5461
11	1	.932313	1.04230	1.33825	1.92272	.173092	.0437515	.101321
12	1	.898510	.966044	1.19077	1.63996	.158724	.0375348	.0844429
13	1	.868704	.901328	1.07048	1.41863	.146682	.0326431	.0715493
14	1	.842156	.845667	.970710	1.24182	.136441	.0287168	.0614726
15	1	.818305	.797246	.886771	1.09810	.127623	.0255118	.0534434

TABLE I (Cont.)

The entries in this table are the moments about the origin $\mu_1'(k,N)$ and the central moments $\mu_1(k,N)$ of the k^{th} order statistic of a sample of size N from a standardized gamma distribution with parameter r.

r = 4

N	k	$\mu_1'(k,N)$	$\mu_2'(k,N)$	$\mu_3'(k,N)$	$\mu_4'(k,N)$	$\mu_2(k,N)$	$\mu_3(k,N)$	$\mu_4(k,N)$
1	1	4.00000	20.0000	120.000	840.000	4.00000	8.00000	72.0000
2	1	2.90625	10.1563	41.2500	190.313	1.70996	1.79425	11.4576
	2	5.09375	29.8438	198.750	1489.69	3.89746	7.02802	6.65339
3	1	2.44942	7.08801	23.4849	87.2606	1.08834	.791765	4.32906
	2	3.81990	16.2927	76.7802	396.416	1.70106	1.54775	10.9225
	3	5.73067	36.6193	259.735	2036.32	3.77865	6.577352	62.5762
4	1	2.18264	5.56858	16.1205	51.9221	.804682	.453680	2.26596
	2	3.24979	11.6463	45.5781	193.276	1.08519	.676751	4.17681
	3	4.39002	20.9391	10.7982	59.9556	1.66687	1.42388	10.3860
	4	6.17756	41.8460	310.319	2515.25	3.68375	6.30009	59.7338
5	1	2.00200	4.65100	12.1841	35.3193	.642989	.298292	1.40379
	2	2.90517	9.23889	31.8660	118.333	.798879	.383691	2.18511
	3	3.76671	15.2574	66.1462	305.691	1.06932	.620217	4.01446
	4	4.80556	24.7270	135.873	795.466	1.63355	1.34595	9.94278
	5	6.52056	46.1257	353.931	2945.19	3.60806	6.11216	57.5834
6	1	1.86896	4.03156	9.76129	26.0465	.538552	.213407	.962974
	2	2.66721	7.74819	24.2983	81.6834	.634167	.249321	1.34524
	3	3.38109	12.2203	47.0014	191.632	.788559	.351150	2.11076
	4	4.15234	18.2946	85.2909	419.750	1.05266	.583759	3.87225
	5	5.13217	27.9432	161.164	983.324	1.60397	1.29138	9.57696
	6	6.79823	49.7622	392.484	3337.56	3.54624	5.97225	55.8852
7	1	1.76548	3.58241	8.12985	20.2686	.465469	.161616	.706918
	2	2.48981	6.72649	19.5499	60.7141	.527356	.176229	.914720
	3	3.11073	10.3024	36.1692	134.107	.625808	.227775	1.30198
	4	3.74157	14.7774	61.4444	268.333	.778150	.330407	2.04528
	5	4.46041	20.9324	103.176	533.313	1.03709	.557196	.375124
	6	5.40087	30.7475	184.359	1163.33	1.57803	1.25078	9.26887
	7	7.03113	52.9314	427.171	3699.94	3.49460	5.86217	54.5023
8	1	1.68188	3.24011	6.96080	16.3862	.411395	.127494	.544528
	2	2.35072	5.97846	16.3132	47.4450	.452551	.131857	.665001
	3	2.90705	8.97060	29.2601	100.521	.519646	.160792	.883769
	4	3.45018	12.5221	47.6845	190.083	.618382	.213808	1.26632
	5	4.03295	17.0328	75.2042	346.583	.768112	.316134	1.96410
	6	4.71689	23.2721	119.959	645.351	1.02307	.535565	3.66371
	7	5.62887	33.2393	205.826	1335.99	1.55506	1.22032	8.99674
	8	7.23145	55.7445	458.792	4037.64	3.45064	5.77198	53.3499
9	1	1.61240	2.96953	6.08401	13.6298	.369713	.103706	.434724
	2	2.23775	5.40475	13.9751	38.4377	.397237	.102820	.507138
	3	2.74614	7.98643	24.4965	78.9706	.445119	.119906	.641409
	4	3.22887	10.9390	38.7872	143.622	.513356	.151655	.858375
	5	3.72682	14.5011	58.8061	248.158	.611917	.202071	1.24205
	6	4.27787	19.0582	88.3227	425.324	.758037	.308645	1.91327
	7	4.93639	25.3790	135.777	755.364	1.01103	.514247	3.59085
	8	5.82672	35.4850	225.841	1501.88	1.53435	1.19877	8.73684
	9	7.40704	58.2769	487.911	4354.61	3.41270	5.69519	52.3790
10	1	1.55238	2.74956	5.40312	11.5888	.336559	.0863914	.356767
	2	2.14350	4.94928	12.2120	31.9981	.354689	.0826928	.400947
	3	2.61474	7.22664	21.0275	64.1961	.389777	.0934143	.487226
	4	3.05276	9.75927	32.5908	113.445	.439936	.112096	.626206
	5	3.49303	12.7085	48.0817	188.889	.507188	.147435	.828098
	6	3.96061	16.2938	69.5305	307.427	.607356	.186261	1.24883
	7	4.48937	20.9012	100.851	503.922	.746708	.313284	1.79662
	8	5.12800	27.2982	150.745	863.126	1.00180	.485095	3.61146
	9	6.00139	37.5318	244.614	1661.57	1.51512	1.18614	8.47766
	10	7.56322	60.5820	514.944	4653.84	3.37964	5.62695	51.5622
11	1	1.50240	2.56672	4.85966	10.0270	.309527	.0733498	.299261
12	1	1.45773	2.41201	4.41615	8.79951	.287040	.0632520	.255511
13	1	1.41813	2.27912	4.04754	7.81345	.268022	.0552528	.221374
14	1	1.38271	2.16357	3.73645	7.00667	.251713	.0487935	.194167
15	1	1.35072	2.06202	3.47047	6.33623	.237561	.0434926	.172089

TABLE I (Cont.)

The entries in this table are the moments about the origin $\mu_1'(k,N)$ and the central moments $\mu_1(k,N)$ of the k^{th} order statistic of a sample of size N from a standardized gamma distribution with parameter r.

r = 5

N	k	$\mu_1'(k,N)$	$\mu_2'(k,N)$	$\mu_3'(k,N)$	$\mu_4'(k,N)$	$\mu_2(k,N)$	$\mu_3(k,N)$	$\mu_4(k,N)$
1	1	5.00000	30.0000	210.000	1680.00	5.00000	10.0000	105.000
2	1	3.76953	16.4648	81.4160	448.301	2.25548	2.34708	18.7116
	2	6.23047	43.5352	338.584	2991.70	4.71642	8.56860	92.7473
3	1	3.24473	12.0106	49.6483	225.793	1.48227	1.05810	7.57900
	2	4.81913	25.3734	144.951	893.317	2.14942	1.95712	16.7291
	3	6.93614	52.6160	435.400	3920.89	4.50599	7.93966	85.3302
4	1	2.93419	9.73073	35.7451	143.414	1.12125	.613342	4.17220
	2	4.17637	18.8501	91.3581	472.928	1.40807	.871979	6.77891
	3	5.46188	31.8967	198.545	1313.71	2.06450	1.77649	15.4044
	4	7.42756	59.5225	514.352	4789.95	4.35386	7.56833	80.3385
5	1	2.72188	8.32061	28.0181	102.534	.911977	.405833	2.68838
	2	3.78343	15.3712	66.6529	306.936	1.05684	.500110	3.69917
	3	4.76577	24.0685	128.416	721.915	1.35593	.787172	6.27053
	4	5.92596	37.1155	245.297	1708.23	1.99845	1.66645	14.4477
	5	7.80290	65.1242	581.616	5560.38	4.23809	7.31509	76.6954
6	1	2.56430	7.35062	23.1151	78.7074	.774999	.291259	1.90455
	2	3.50979	13.1706	52.5333	221.665	.851901	.327348	2.35473
	3	4.33071	19.7725	94.8923	477.478	1.01745	.450905	3.42275
	4	5.20083	28.3645	161.940	966.353	1.31587	.733884	5.90474
	5	6.28853	41.4910	286.976	2079.17	1.94538	1.59097	13.7107
	6	8.10584	69.8509	640.544	6256.62	4.54619	7.12700	73.8885
7	1	2.44095	6.63632	19.7300	63.3342	.678084	.220795	1.43679
	2	3.30439	11.6364	43.4254	170.947	.717470	.232489	1.64580
	3	4.02331	17.0058	75.3031	348.462	.818796	.294647	2.17577
	4	4.74057	23.4613	121.011	649.501	.988333	.420706	3.22340
	5	5.54602	32.0418	192.636	1203.99	1.28348	.695188	5.62838
	6	6.58554	45.2707	32.4711	2429.25	1.90137	1.53641	13.1096
	7	8.35923	73.9476	693.183	6894.52	4.07088	6.97923	71.6417
8	1	2.34074	6.08476	17.2525	52.6953	.605699	.174092	1.13318
	2	3.14242	10.4972	37.0729	137.807	.622424	.174396	1.22528
	3	3.79028	15.0541	62.4827	270.366	.687812	.209473	1.51318
	4	4.41170	20.2588	96.6703	478.621	.795750	.273397	2.05897
	5	5.06944	26.6638	145.352	820.381	.964621	.401367	3.05365
	6	5.83196	35.2686	221.006	1434.16	1.25581	.662053	5.43480
	7	6.83672	48.6047	359.280	2760.94	1.86392	1.49640	12.5919
	8	8.57673	77.5080	740.883	7485.03	4.00770	6.85752	69.8041
9	1	2.25706	5.64375	15.3599	44.9463	.549445	.141376	.923734
	2	3.01020	9.61284	32.3933	114.687	.551533	.136278	.953339
	3	3.60518	13.5925	53.4514	218.725	.595209	.156166	1.12606
	4	4.16049	17.9772	80.5454	373.648	.667474	.197501	1.41338
	5	4.72569	23.1109	116.826	609.837	.778670	.252076	2.00847
	6	5.34445	29.5062	168.173	988.816	.943062	.397960	2.83433
	7	6.07572	38.1498	247.423	1656.83	1.23533	.624974	5.35809
	8	7.05413	51.5918	391.239	3076.40	1.83102	1.47068	12.0925
	9	8.76705	80.8150	784.589	8036.11	3.95380	6.75301	68.2812
10	1	2.18568	5.28158	13.8661	39.0781	.504372	.117478	.772331
	2	2.89242	8.90324	28.8036	97.7597	.496695	.109602	.767847
	3	3.45334	12.4513	46.7521	182.397	.525692	.122755	.867138
	4	3.95947	16.2555	69.0830	303.490	.578065	.142452	1.07808
	5	4.46202	20.5597	97.7388	478.885	.650038	.199998	1.26874
	6	4.98935	25.6620	13.5914	740.790	.768412	.209406	2.14685
	7	5.58115	32.0689	189.678	1154.17	.919651	.430961	2.38894
	8	6.28773	40.7560	272.171	1872.26	1.22045	.560151	5.59678
	9	7.24574	54.3008	421.006	3377.44	1.80013	1.46951	11.4765
	10	8.93609	83.7610	824.987	8553.74	3.90734	6.65833	67.0331
11	1	2.12378	4.97782	12.6565	34.4962	.467381	.0994281	.658834
12	1	2.06936	4.71866	11.6563	30.8291	.436426	.0854204	.571221
13	1	2.02097	4.49441	10.8150	27.8341	.410104	.0743076	.501930
14	1	1.97753	4.29804	10.0971	25.3460	.387418	.0653212	.446025
15	1	1.93823	4.12436	9.47703	23.2490	.367640	.0579414	.400129

TABLE II

The two entries in each cell of this table are the modal values of the smallest (top) and the largest (bottom) of a sample of size N from the standardized gamma distribution with parameter r.

N \ r	1	2	3	4	5	6	7	8	9	10
1	0.0000 0.0000	1.0000 1.0000	2.0000 2.0000	3.0000 3.0000	4.0000 4.0000	5.0000 5.0000	6.0000 6.0000	7.0000 7.0000	8.0000 8.0000	9.0000 9.0000
2	0.0000 0.6932	0.7070 1.9375	2.0000 3.1143	2.3724 4.2594	3.2442 5.3855	4.1300 6.4983	5.0260 7.6014	5.9298 8.6970	6.8399 9.7863	7.7552 10.8706
3	0.0000 1.0986	0.5773 2.4647	1.3007 3.7301	2.0774 4.9486	2.8849 6.1384	3.7131 7.3085	4.5564 8.4640	5.4115 9.6081	6.2760 10.7429	7.1482 11.8701
4	0.0000 1.3863	0.5000 2.8310	1.1654 4.1542	1.8942 5.4204	2.6599 6.6520	3.4506 7.8595	4.2596 9.0493	5.0838 10.2252	5.9174 11.3899	6.7615 12.5452
5	0.0000 1.6094	0.4472 3.1113	1.0710 4.4767	1.7650 5.7779	2.5003 7.0400	3.2636 8.2750	4.0474 9.4900	4.8473 10.6892	5.6601 11.8758	6.4835 13.0518
6	0.0000 1.7918	0.4083 3.3382	1.0000 4.7365	1.6671 6.0650	2.3787 7.3510	3.1205 8.6075	3.8848 9.8422	4.6664 11.0597	5.4621 12.2635	6.2694 13.4557
7	0.0000 1.9459	0.3780 3.5285	0.9440 4.9536	1.5892 6.3045	2.2815 7.6100	3.0059 8.8841	3.7542 10.1349	4.5209 11.3674	5.3026 12.5852	6.0967 13.7906
8	0.0000 2.0794	0.3535 3.6923	0.8982 5.1400	1.5251 6.5096	2.2012 7.8316	2.9110 9.1205	3.6458 10.3849	4.4001 11.6300	5.1700 12.8596	5.9529 14.0763
9	0.0000 2.1972	0.3333 3.8361	0.8597 5.3032	1.4710 6.6889	2.1333 8.0250	2.8306 9.3267	3.5538 10.6028	4.2972 11.8588	5.0570 13.0986	5.8304 14.3249
10	0.0000 2.3026	0.3162 3.9642	0.8269 5.4482	1.4245 6.8481	2.0747 8.1966	2.7610 9.5094	3.4741 10.7958	4.2081 12.0613	4.9590 13.3100	5.7240 14.5448
15	0.0000 2.7081	0.2582 4.4525	0.7125 5.9986	1.2607 7.4504	1.8668 8.8445	2.5129 10.1984	3.1889 11.5224	3.8883 12.8230	4.6066 14.1047	5.3408 15.3705
20	.0000 2.9957	0.2236 4.7950	0.6416 6.3824	1.1576 7.8689	1.7345 9.2933	2.3539 10.6747	3.0053 12.0240	3.6816 13.3481	4.3782 14.6518	5.0918 15.9385
25	0.000 3.4012	0.2000 5.0587	0.5919 6.6766	1.0841 8.1889	1.6395 9.6360	2.2393 11.0378	2.8724 12.4059	3.5316 13.7476	4.2121 15.0677	4.9103 16.3700
50	0.0000 3.9120	0.1414 5.8678	0.4618 7.5740	0.8874 9.1606	1.3818 10.6733	1.9251 12.1346	2.5056 13.5573	3.1155 14.9497	3.7494 16.3177	4.4033 17.6652
100	0.0000 4.6052	0.1000 6.6643	0.3615 8.4501	0.7297 10.1040	1.1703 11.6763	1.6635 13.1915	2.1969 14.6639	2.7624 16.1027	3.3542 17.5143	3.9681 18.9031

3. Covariance Between the mth and nth Order Statistics

Let $\xi = x_{[m]}$ and $\eta = x_{[n]}$ be the mth and nth ($n > m$) order statistics from the gamma distribution (1.1). Then, we have

$$E(\xi\eta) = C \int_0^\infty \int_\xi^\infty \xi\eta G_r^{m-1}(\xi)[G_r(\eta) - G_r(\xi)]^{n-m-1} \cdot [1 - G_r(\eta)]^{N-n} g_r(\xi) g_r(\eta)\, d\xi\, d\eta \tag{3.1}$$

where

$$C = \frac{N!}{(m-1)!\,(n-m-1)!\,(N-n)!}.$$

The expression on the right side of (3.1) can be written as

$$E(\xi\eta) = C \sum_{\alpha,\beta} (-1)^{\alpha+\beta} \binom{m-1}{\alpha} \binom{n-m-1}{\beta} \cdot \int_0^\infty e^{-qu} \left[\sum_{j=0}^{r-1} \frac{u^j}{j!}\right]^{q-1} \frac{u^r}{\Gamma(r)}\, du \int_0^\infty e^{-sv} \left[\sum_{j=0}^{r-1} \frac{v^j}{j!}\right]^{s-1} \frac{v^r}{\Gamma(r)}\, dv \tag{3.2}$$

where $q = \alpha + n - m - \beta$ and $s = N - n + \beta + 1$ and the first summation on the right side is over the positive integers α, β ($0 \leqq \alpha \leqq m - 1$, $0 \leqq \beta \leqq n - m - 1$). As before, let $a_t(r, p)$ denote the coefficient of u^t in the expansion for

$$\left[\sum_{j=0}^{r-1} \frac{u^j}{j!}\right]^p,$$

then (3.2) simplifies to

$$E(\xi\eta) = \frac{C}{(\Gamma(r))^2} \sum_{\alpha,\beta} (-1)^{\alpha+\beta} \binom{m-1}{\alpha} \binom{n-m-1}{\beta} \cdot \sum_{t=0}^{(r-1)(s-1)} \frac{a_t(r, s-1)\Gamma(t+r+1)}{s^{t+r+1}} \cdot \left[\sum_{\alpha,\beta}^{t+r} \binom{s^j}{j!} \sum_{p=0}^{(r-1)(q-1)} \frac{\Gamma(j+r+p+1)a_p(r, q-1)}{(s+q)^{j+r+p+1}}\right] \tag{3.3}$$

or

$$E(\xi\eta) = \frac{C}{(\Gamma(r))^2} \sum_{\alpha,\beta,t,j,p} (-1)^{\alpha+\beta} \binom{m-1}{\alpha} \binom{n-m-1}{\beta} \cdot \left[\frac{a_t(r, s-1)a_p(r, q-1)\Gamma(t+r+1)\Gamma(j+r+p+1)}{j!\, s^{t+r-j+1}(s+q)^{j+r+p+1}}\right] \tag{3.4}$$

From (3.4) we can obtain the covariance (ξ, η) by subtracting from $E(\xi\eta)$ the product $E(\xi)E(\eta)$ which we get from the results of Section 2.

4. Some Remarks about the Distribution of the kth Order Statistic

a. The Mode of the kth Order Statistic

The modal value t of the kth order statistic from a sample of size N from the gamma distribution with parameter r is given by

$$\frac{\partial}{\partial t}\left[\bar{e}^{Nt}t^{r-1}\psi(r, t)\{e^t - \psi(r, t)\}^{N-k}\{e^t - \psi(r, t)\}^{k-1}\right] = 0 \tag{4.1}$$

where

$$\psi(r, t) = \sum_{j=0}^{r-1}\frac{t^j}{j!}, \qquad r \geqq 1$$
$$= 0, \quad \text{otherwise.} \tag{4.2}$$

The function $\psi(r, t)$ satisfies the following difference-differential equation

$$\frac{\partial\psi(r, t)}{\partial t} = \psi'(r, t) = \psi(r-1, t). \tag{4.3}$$

Using equation (4.3) in (4.1) we obtain after some simplification

$$(-Nt + r - 1)\psi(r, t)[e^t - \psi(r, t)] + (N-k)t\psi(r-1, t)[e^t - \psi(r, t)] + (k-1)t\psi(r, t)[e^t - \psi(r-1, t)] = 0 \tag{4.4}$$

or

$$(r - t - 1)\left(\sum_{j=0}^{r-1}\frac{t^j}{j!}\right)\left[e^t - \sum_{j=0}^{r-1}\frac{t^j}{j!}\right] = \frac{t^r}{(r-1)!}\left[(N-k)e^t - (N-1)\sum_{j=0}^{r-1}\frac{t^j}{j!}\right] \tag{4.5}$$

Equation (4.5) can be used for an iterative procedure to solve for t. The cases $k = 1$ and $k = N$ are of particular interest and we obtain

$$\sum_{j=0}^{r-1}\frac{\bar{e}^t t^j}{j!} = \frac{N-1}{(r-1)!}\frac{\bar{e}^t t^r}{(r-t-1)} \tag{4.6}$$

$$\sum_{j=0}^{r-1}\frac{\bar{e}^t t^j}{j!} = \frac{N-1}{(r-1)!}\frac{\bar{e}^t t^r}{(r-t-1)} + 1 \tag{4.7}$$

which give the modal values of the smallest and the largest order statistics, respectively. It should be noted that the present forms (4.6) and (4.7) are helpful in that the value of the left side expression can be obtained from a table of cumulative Poisson distribution.

It is easy to show that there are unique solutions for (4.6) and (4.7) in the open intervals $0 < t < r - 1$ and $r - 1 < t < \infty$, respectively, provided $N \geqq 2$. For $N = 1$, $t = r - 1$, since in this case the mode is the same as that of a gamma distribution. Also, t increases monotonically with N for the largest order statistic and decreases monotonically with N for the smallest order statistic.

Another particular case of interest is $r = 1$, since in this case the gamma distribution reduces to the exponential distribution. For this case equation (4.5)

reduces to the simple and explicit form

$$t = \log_e N - \log_e (N - k + 1). \tag{4.8}$$

Equations (4.6) and (4.7) were solved by iterative methods on IBM 650 for selected values of N and r. The results of these computations are given in Table II.

b. Cumulative Distribution Function of the Order Statistics

The cumulative distribution function $F_r(y; k, N)$ of the kth order statistic $y = x_{(k)}$ from the gamma distribution is

$$\begin{aligned} F_r(y; k, N) &= \frac{N!}{(k-1)!\,(N-k)!} \int_0^y G_r^{k-1}(x)[1 - G_r(x)]^{N-k} g_r(x)\, dx \\ &= \frac{N!}{(k-1)!\,(N-k)!} \int_0^{G_r(y)} u^{k-1}(1-u)^{N-k}\, du \\ &= I_{G_r(y)}(k, N-k+1), \end{aligned} \tag{4.9}$$

where $I_{G_r(y)}(k, N-k+1)$ is the ratio of the incomplete beta function to the complete beta function and is tabulated in [5]. For a given value of y, we can compute the c.d.f. $F_r(y; k, N)$ by first consulting a table of the cumulative distribution of the Poisson or gamma distribution to get $G_r(y)$ and then consulting the table of the cumulative of the beta distribution. Again,

$$F_r(y; k, N) = \frac{N!}{(k-1)!\,(N-k)!} \sum_{\alpha=0}^{N-k} \frac{(-1)^\alpha}{\alpha + k} \binom{N-k}{\alpha} \left\{ \sum_{j=r}^{\infty} \frac{\bar{e}^y y^j}{j!} \right\}^{\alpha+k}, \tag{4.10}$$

$$F_r(y; k, N) = \frac{N!}{(k-1)!\,(N-k)!} \left[\sum_{j=k}^{N} F_r(y; j, j) \binom{N-k}{j-k} \frac{(-1)^{j-k}}{j} \right]. \tag{4.11}$$

The function $F_r(y; j, j)$ is the c.d.f. of the largest of jx's and using K. Pearson's notation [6] for the incomplete gamma function can be written as

$$F_r(y; j, j) = G_r^j(y) = I^j\left(\frac{y}{\sqrt{r}}, r - 1\right) = \left\{ \sum_{j=r}^{\infty} \frac{\bar{e}^y y^j}{j!} \right\}^j. \tag{4.12}$$

Also for $k = 1$ and $k = N$ i.e. for the distribution of the smallest and the largest respectively, we get the following simple forms

$$F_r(y; 1, N) = 1 - [1 - G_r(y)]^N = 1 - \bar{e}^{Ny} \left[\sum_{j=0}^{r-1} \frac{y^j}{j!} \right]^N \tag{4.13}$$

$$F_r(y; N, N) = G^N(y) = \bar{e}^{Ny} \left[\sum_{j=r}^{\infty} \frac{y^j}{j!} \right]^N. \tag{4.14}$$

The c.d.f. of the median i.e. of $x_{(m+1)}$ when $N = 2m + 1$, is

$$F_r(y; m+1, 2m+1) = \frac{(2m+1)!}{(m!)^2} \sum_{\alpha=0}^{m} \frac{(-1)^\alpha}{\alpha + m + 1} \binom{m}{\alpha} \left\{ \sum_{j=r}^{\infty} \frac{\bar{e}^y y^j}{j!} \right\}^{\alpha+m+1}. \tag{4.15}$$

Probabilities in (4.10) ... (4.15) can be obtained with the help of a table of cumulative Poisson distribution [3] and also with the help of a table of incomplete gamma function [6].

TABLE III A

r = 1

This table gives the α-percentage points of the k^{th} smallest order statistic of size N [$1 \le k \le N$, $N = 1(1)10, 25, 50, 100, 200, 500, 1000$, and selected values of α] from a standardization gamma distribution with parameter r.

N	k \ α	.01	.05	.10	.25	.50	.75	.90	.95	.99
1	1	.01005	.05129	.10536	.28768	.69315	1.38629	2.30259	2.99573	4.60517
2	1	.00503	.02565	.05268	.14384	.34657	.69315	1.15129	1.49785	2.30259
	2	.10536	.25310	.38013	.69315	1.22796	2.01014	2.96973	3.67612	5.29580
3	1	.00335	.01710	.03512	.09589	.23105	.46210	.76753	.99859	1.53507
	2	.06071	.14543	.21791	.39504	.69315	1.11978	1.63066	1.99989	2.83186
	3	.24263	.45950	.62392	.99414	1.57842	2.39207	3.36648	4.07737	5.70044
4	1	.00251	.01282	.02634	.07192	.17328	.34657	.57565	.74893	1.15129
	2	.04291	.10271	.15380	.27842	.48732	.78456	1.13800	1.39191	1.95992
	3	.15184	.28582	.38634	.60939	.95262	1.41461	1.94799	2.32677	3.17011
	4	.38013	.64031	.82631	1.22796	1.83822	2.66794	3.64981	4.36293	5.98771
5	1	.00201	.01026	.02107	.05754	.13863	.27726	.46051	.59915	.94209
	2	.03323	.07952	.11905	.21537	.37660	.60547	.87681	1.07122	1.50476
	3	.11165	.20981	.28321	.44541	.69315	1.02321	1.39983	1.66463	2.24772
	4	.25112	.41945	.53804	.78926	1.15897	1.64114	2.18712	2.57125	3.42093
	5	.50768	.79691	.99685	1.41824	2.04446	2.88396	3.87031	4.58478	6.21062
6	1	.00168	.00855	.01756	.04795	.11552	.23105	.36920	.49929	.76753
	2	.02713	.06491	.09717	.17574	.30714	.49344	.71400	.87180	1.22312
	3	.08854	.16624	.22428	.35228	.54716	.80564	1.09904	1.30438	1.75406
	4	.19004	.31655	.40525	.59204	.86415	1.21429	1.60490	1.87627	2.46829
	5	.34858	.54163	.67272	.94294	1.33010	1.82536	2.37952	2.76700	3.62073
	6	.62392	.93384	1.14347	1.57842	2.21549	3.06153	4.05088	4.76623	6.39272
7	1	.00144	.00733	.01505	.04110	.09902	.19804	.32893	.42796	.65788
	2	.02293	.05485	.08210	.14847	.25999	.41659	.60250	.73543	1.03103
	3	.07344	.13784	.18590	.29178	.45275	.66573	.90679	1.07511	1.44257
	4	.15347	.25531	.32656	.47618	.69315	.97061	1.27798	1.49023	1.95003
	5	.26961	.41743	.51721	.72134	1.01027	1.37409	1.77408	2.04981	2.64784
	6	.44105	.65258	.79283	1.07672	1.47626	1.98072	2.54055	2.93041	3.78693
	7	.72971	1.05509	1.27186	1.71669	2.36149	3.21230	4.20378	4.91977	6.54659
8	1	.00126	.00641	.01317	.03596	.08664	.17328	.28783	.37446	.57565
	2	.01985	.04708	.07109	.12855	.22456	.36054	.52130	.63616	.89145
	3	.06277	.11778	.15882	.24921	.38643	.56777	.77267	.91552	1.22680
	4	.12891	.21431	.27399	.39911	.58010	.81078	1.06531	1.24050	1.66495
	5	.22090	.34142	.42254	.58790	.82062	1.11145	1.42853	1.64558	2.11238
	6	.34705	.51134	.61949	.83653	1.13781	1.51149	1.91834	2.19723	2.79951
	7	.52773	.75358	.90079	1.19501	1.60380	2.11503	2.67917	3.07110	3.92410
	8	.82631	1.16366	1.38585	1.83822	2.48891	3.34308	4.33638	5.05283	6.67943
9	1	.00112	.00570	.01171	.03196	.07702	.15404	.25585	.33286	.51169
	2	.01744	.04187	.06269	.11334	.19799	.31783	.45944	.56061	.78533
	3	.05485	.10285	.13868	.21750	.33580	.49518	.67350	.79771	1.06802
	4	.11122	.18482	.23623	.34389	.49936	.69714	.91484	1.08438	1.38617
	5	.18759	.28951	.35806	.49751	.69315	.93660	1.20074	1.38083	1.76827
	6	.28772	.42303	.51179	.68918	.93374	1.23447	1.55874	1.77934	2.25132
	7	.42112	.59849	.71319	.94022	1.25092	1.63234	2.04414	2.32550	2.93043
	8	.60875	.84597	.99869	1.30096	1.71691	2.23348	2.80081	3.19402	4.05745
	9	.91499	1.26185	1.48837	1.94659	2.60194	3.45907	4.45348	5.17026	6.80331
10	1	.00101	.00513	.01054	.02877	.06931	.13863	.23026	.29958	.46051
	2	.01524	.03700	.05604	.10136	.17705	.28418	.41075	.50114	.70189
	3	.05104	.09123	.12737	.19303	.29918	.43919	.59711	.70704	.94608
	4	.09806	.16255	.21042	.30225	.43866	.61194	.80236	.93298	1.21362
	5	.16304	.25159	.31105	.43183	.60091	.81078	1.03784	1.19224	1.52170
	6	.24634	.36174	.43730	.58790	.79476	1.04788	1.31931	1.50310	1.89419
	7	.35257	.49385	.59470	.78154	1.03536	1.34382	1.66201	1.89693	2.37082
	8	.49145	.67944	.79949	1.03440	1.35259	1.73987	2.12370	2.43903	3.00051
	9	.68448	.93100	1.08812	1.39687	1.81856	2.33935	2.90958	3.31536	4.19147
	10	.99685	1.35143	1.58148	2.04446	2.70351	3.56285	4.55819	5.27534	6.90327
25	25	1.78239	2.18100	2.43054	2.91984	3.59922	4.47053	5.47136	6.19011	7.81931
50	50	2.43054	2.84464	3.10093	3.59922	4.28546	5.16080	6.16344	6.88277	8.51221
100	100	3.10093	3.52292	3.78263	4.28546	4.97515	5.85250	6.85611	7.57562	9.20535
200	200	3.78263	4.20861	4.47004	4.97515	5.66657	6.54492	7.54907	8.26877	9.89950
500	500	4.69204	5.12042	5.38287	5.88938	6.58184	7.46079	8.46508	9.18467	10.81479
1000	1000	5.38287	5.81208	6.07489	6.58184	7.27462	8.15394	9.15870	9.87782	11.51293

For given N, k, ($1 \le k \le N$) and α the entries in the above table are the values of y for which $\frac{N!}{(k-1)!(N-k)!} \int_0^y G_r^{k-1}(z) \left[1-G_r(z)\right]^{N-k} g_r(z)dz = \alpha$, where $g_r(z)$ and $G_r(z)$ refer to the p.d.f. and c.d.f. of the standardized gamma chance variable, respectively.

TABLE III B

r = 2

This table gives the α-percentage points of the k^{th} smallest order statistic of a sample of size N [1 ≤ k ≤ N, N = 1(1)10, 25, 50, 100, 200, 500, 1000, and selected values of α] from a standardized gamma distribution with parameter r.

N	k \ α	.01	.05	.10	.25	.50	.75	.90	.95	.99
1	1	.14855	.35536	.53181	.96128	1.67835	2.69263	3.88972	4.74386	6.63835
2	1	.10363	.24389	.36062	.63623	1.07795	1.67835	2.36459	2.84449	3.88972
	2	.53181	.88931	1.14176	1.67835	2.47297	3.51827	4.71237	5.55659	7.42727
3	1	.08410	.19649	.28894	.50407	.84178	1.29116	1.79556	2.14412	2.89469
	2	.39005	.64032	.81276	1.16958	1.67835	2.31964	3.02259	3.50510	4.54471
	3	.86692	1.28654	1.56644	2.13760	2.95286	4.00189	5.18930	6.02715	7.88478
4	1	.07257	.16880	.24741	.42865	.70946	1.07795	1.48654	1.76653	2.36459
	2	.32221	.52414	.66156	.94228	1.33558	1.82201	2.34566	2.70033	3.45364
	3	.65661	.95747	1.15338	1.54257	2.07637	2.73139	3.43825	3.92009	4.95421
	4	1.14176	1.59320	1.88628	2.47297	3.29588	4.34390	5.52554	6.35883	8.20776
5	1	.06475	.15016	.21957	.37862	.62276	.93989	1.28834	1.52564	2.06075
	2	.28040	.45348	.57036	.80709	1.13513	1.53609	1.96301	2.24982	2.85383
	3	.54972	.79458	.95214	1.26139	1.67835	2.18013	2.71117	3.06772	3.88962
	4	.88510	1.21445	1.42308	1.82929	2.37552	3.03652	3.74400	4.22463	5.25418
	5	1.37099	1.84110	2.14157	2.73635	3.56228	4.60819	5.78502	6.61488	8.45742
6	1	.05900	.13652	.19929	.34243	.56058	.84178	1.12116	1.35645	1.79556
	2	.25135	.40482	.50785	.71536	1.00057	1.34627	1.71151	1.95544	2.46617
	3	.48178	.69238	.82691	1.08894	1.43858	1.85456	2.28990	2.57965	3.18574
	4	.74919	1.01930	1.18844	1.51385	1.94388	2.45375	2.98825	3.34539	4.09692
	5	1.08182	1.42918	1.64562	2.06202	2.61527	3.27910	3.98619	4.46545	5.49118
	6	1.56644	2.04849	2.35347	2.95286	3.77991	4.82342	5.99626	6.82334	8.66079
7	1	.05455	.12599	.18368	.31474	.51334	.76774	1.04369	1.22922	1.62170
	2	.22968	.36875	.46171	.64807	.90389	1.20924	1.53119	1.74533	2.19153
	3	.43368	.62069	.73949	.96963	1.27451	1.63437	2.00809	2.25544	2.76951
	4	.66071	.89399	1.03904	1.31604	1.67835	2.10298	2.54293	2.83419	3.44088
	5	.92403	1.21077	1.38743	1.72312	2.16123	2.67588	3.21202	3.56913	4.31915
	6	1.25356	1.61313	1.83481	2.25777	2.81528	3.48044	4.18664	4.66466	5.68712
	7	1.73632	2.22649	2.53440	3.13655	3.96362	5.00487	6.17428	6.99910	8.83226
8	1	.05096	.11755	.17119	.29267	.47594	.70946	.96158	1.13110	1.48654
	2	.21268	.33902	.42587	.59612	.82752	1.10472	1.39445	1.58645	1.98508
	3	.39734	.56685	.67409	.88102	1.15354	1.47333	1.80355	2.02116	2.47134
	4	.59711	.80470	.93313	1.17710	1.49399	1.86247	2.24128	2.49056	3.06814
	5	.81941	1.06800	1.22009	1.50699	1.87756	2.30771	2.75040	3.04243	3.64904
	6	1.07888	1.37731	1.55917	1.90190	2.34540	2.86292	3.39993	3.75681	4.50520
	7	1.40549	1.77379	1.99910	2.42657	2.98679	3.65241	4.35771	4.83497	5.84806
	8	1.88628	2.38218	2.69203	3.29588	4.12257	5.16138	6.32810	7.15097	8.98001
9	1	.04801	.11060	.16091	.27459	.44541	.66215	.89514	1.05137	1.37792
	2	.19859	.31794	.39704	.55449	.76762	1.02183	1.28643	1.46130	1.82320
	3	.36875	.52454	.62288	.81185	1.05710	1.34930	1.64688	1.84233	2.24519
	4	.54853	.73696	.85308	1.07277	1.35656	1.68471	2.02015	2.23994	2.69247
	5	.74325	.96501	1.10000	1.35336	1.67835	2.05258	2.43466	2.68514	3.20177
	6	.96136	1.22097	1.37809	1.67200	2.04834	2.48210	2.92649	3.21891	3.82517
	7	1.21750	1.52456	1.71021	2.05798	2.50516	3.02482	3.56187	3.91850	4.66468
	8	1.54151	1.91628	2.14428	2.57490	3.13685	3.80265	4.50679	4.98293	6.00392
	9	2.02037	2.52043	2.83167	3.43644	4.26255	5.29955	6.46350	7.28466	9.11758
10	1	.04551	.10473	.15227	.25942	.41990	.62276	.84005	.98540	1.28834
	2	.18489	.29722	.37314	.52020	.71851	.95412	1.19855	1.35965	1.69227
	3	.35440	.49002	.59296	.75616	.98460	1.25021	1.52227	1.70047	2.06667
	4	.51050	.68332	.79596	.99078	1.24926	1.54676	1.84950	2.04720	2.45281
	5	.68451	.88611	1.00837	1.23694	1.52857	1.86247	2.20144	2.42266	2.87668
	6	.87488	1.10702	1.24681	1.50699	1.83778	2.21603	2.60034	2.85158	3.36864
	7	1.08949	1.35742	1.51827	1.81733	2.19783	2.63421	3.06423	3.37220	3.97531
	8	1.34279	1.65642	1.84494	2.19643	2.64630	3.16708	3.66346	4.06124	4.74965
	9	1.66450	2.04425	2.27419	2.70713	3.27022	3.93588	4.63935	5.12827	6.15998
	10	2.14157	2.64470	2.95696	3.56228	4.38761	5.42268	6.58427	7.40406	9.22844
25	25	3.22291	3.73625	4.04991	4.65182	5.46574	6.48318	7.62616	8.43448	10.23867
50	50	4.04991	4.56029	4.87094	5.46573	6.26909	7.27391	8.40464	9.20572	10.99685
100	100	4.87094	5.37537	5.68205	6.26909	7.06235	8.05592	9.17616	9.97087	11.75096
200	200	5.68205	6.17990	6.48262	7.06235	7.84661	8.83040	9.94165	10.73101	12.50236
500	500	6.73821	7.22796	7.52599	8.09736	8.87149	9.84445	10.94542	11.72852	13.48812
1000	1000	7.52599	8.01047	8.30549	8.87149	9.63917	10.60541	11.70033	12.47895	14.23663

For given N, k, (1 ≤ k ≤ N) and α, the entries in the above table are the values of y for which $\frac{N!}{(k-1)!(N-k)!}\int_0^y G_r(z)^{k-1}\left[1-G_r(z)\right]^{N-k} g_r(z)dz = \alpha$, where $g_r(z)$ and $G_r(z)$ refer to the p.d.f. and c.d.f. of the standardized gamma chance variable, respectively.

TABLE III C

r = 3

This table gives the α-percentage points of the k^{th} smallest order statistic of a sample of size N [1 ≤ k ≤ N, N = 1(1)10, 25, 50, 100, 200, 500, 1000, and selected values of α] from a standardization gamma distribution with parameter r.

N	k \ α	.01	.05	.10	.25	.50	.75	.90	.95	.99
1	1	.43605	.81769	1.10207	1.72730	2.67406	3.92040	5.32232	6.29579	8.40595
2	1	.33817	.62178	.82656	1.26086	1.88744	2.67406	3.52510	4.10147	5.32232
	2	1.10207	1.62668	1.97379	2.67406	3.65638	4.89287	6.26019	7.20790	9.27067
3	1	.29218	.53226	.70303	1.05886	1.55977	2.17301	2.82228	3.25570	4.16107
	2	.87562	1.26697	1.51846	2.01118	2.67406	3.47044	4.31244	4.87757	6.07034
	3	1.59535	2.16691	2.53118	3.24769	4.23001	5.45120	6.79717	7.73102	9.76927
4	1	.26364	.47752	.62823	.93880	1.36916	1.88744	2.42827	2.78570	3.52510
	2	.76114	1.09016	1.29858	1.70092	2.23154	2.85554	3.50209	3.92961	4.81770
	3	1.29122	1.72202	1.98942	2.50052	3.17223	3.96672	4.79978	5.35725	6.53302
	4	1.97379	2.56548	2.93609	3.65638	4.63359	5.84213	7.17326	8.09787	10.12013
5	1	.24355	.43938	.57644	.85667	1.24067	1.69759	2.16929	2.47873	3.15294
	2	.68787	.97878	1.16134	1.51037	1.96485	2.49218	3.03177	3.38527	4.11256
	3	1.12972	1.49245	1.71492	2.13363	2.67406	3.29993	3.94257	4.36567	5.24275
	4	1.62095	2.07121	2.34598	2.88468	3.53836	4.32888	5.15434	5.70614	6.86989
	5	2.27797	2.87951	3.25257	3.97264	4.94398	6.14229	7.46226	8.38012	10.39075
6	1	.22834	.41070	.53769	.79577	1.14639	1.55977	1.94600	2.25892	2.82228
	2	.63546	.89996	1.06478	1.37779	1.78157	2.24558	2.71614	3.02236	3.64816
	3	1.02380	1.34410	1.53863	1.90237	2.36610	2.89639	3.43420	3.78493	4.50452
	4	1.42702	1.80732	2.03646	2.46354	3.00796	3.63315	4.27187	4.69146	5.56013
	5	1.89270	2.35389	2.63243	3.15450	3.82771	4.61396	5.43319	5.98038	7.13492
	6	2.53118	3.13778	3.51158	4.23001	5.19580	6.38562	7.69678	8.60932	10.61084
7	1	.21627	.38805	.50719	.74820	1.07335	1.45385	1.84074	2.09182	2.60191
	2	.59542	.84020	.99193	1.27852	1.64732	2.06426	2.48588	2.75893	3.31391
	3	.94693	1.23757	1.41295	1.73887	2.15100	2.61808	3.08775	3.39213	4.01221
	4	1.29732	1.65345	1.83437	2.20583	2.67406	3.20505	3.74076	4.08922	4.80284
	5	1.67549	2.06629	2.29947	2.73084	3.27674	3.90036	4.53556	4.95192	5.81394
	6	2.12324	2.59097	2.87161	3.39498	4.06672	4.84889	5.66274	6.20623	7.35334
	7	2.74755	3.35673	3.73049	4.44672	5.40727	6.59003	7.89393	8.80219	10.79615
8	1	.20635	.36953	.48234	.70961	1.01454	1.36916	1.72771	1.95942	2.42827
	2	.56337	.78998	.93428	1.20052	1.53950	1.92377	2.30864	2.55683	3.05919
	3	.88766	1.15598	1.31713	1.61521	1.98964	2.41118	2.83234	3.10394	3.65441
	4	1.20201	1.56695	1.68818	2.02127	2.43790	2.90629	3.37483	3.67762	4.36617
	5	1.52795	1.87391	2.07872	2.45468	2.92518	3.45591	3.98940	4.33585	5.04455
	6	1.88870	2.28624	2.52185	2.95560	3.50177	4.12336	4.75512	5.16914	6.02552
	7	2.32303	2.79486	3.07660	3.60026	4.27014	5.04845	5.85785	6.39866	7.53231
	8	2.93609	3.54646	3.91968	4.63359	5.58949	6.76585	8.06394	8.96857	10.95566
9	1	.19801	.35400	.46152	.67748	.96583	1.29945	1.63506	1.85122	2.28703
	2	.53633	.75376	.88717	1.13704	1.45368	1.81079	2.16676	2.39560	2.85703
	3	.84019	1.09078	1.24086	1.51716	1.85905	2.24955	2.63403	2.88105	3.37961
	4	1.12787	1.40928	1.57577	1.88040	2.25907	2.68214	3.10270	3.37319	3.92021
	5	1.41842	1.73246	1.91737	2.25487	2.67406	3.14284	3.61006	3.91143	4.52334
	6	1.72741	2.07990	2.28726	2.66600	3.13759	3.66741	4.19878	4.54345	5.24799
	7	2.07528	2.47734	2.71450	3.14951	3.69524	4.31507	4.94350	5.35542	6.20626
	8	2.49915	2.97356	3.25590	3.77922	4.44707	5.22203	6.02732	6.56535	7.70527
	9	3.10297	3.71365	4.08622	4.79767	5.74941	6.92069	8.21335	9.11485	11.10404
10	1	.19083	.34070	.44376	.65014	.92459	1.24067	1.55731	1.76066	2.16929
	2	.50957	.71763	.84753	1.08404	1.38239	1.71737	2.04998	2.26313	2.69174
	3	.81608	1.03680	1.19573	1.43712	1.75956	2.11879	2.47439	2.70214	3.16026
	4	1.06891	1.33076	1.49444	1.76808	2.11753	2.50590	2.89005	3.13619	3.63201
	5	1.33252	1.62238	1.79231	2.10115	2.48250	2.90629	3.32606	3.59552	4.13970
	6	1.60657	1.92688	2.11427	2.45468	2.87534	3.34393	3.80978	4.10989	4.71862
	7	1.90311	2.26020	2.46923	2.84966	3.32164	3.85043	4.36157	4.72277	5.42069
	8	2.24101	2.64618	2.88432	3.31993	3.86493	4.48261	5.06123	5.51926	6.30233
	9	2.65646	3.13254	3.41503	3.93773	4.60356	5.37540	6.17758	6.72869	7.87811
	10	3.25257	3.86302	4.23485	4.94398	5.89189	7.05841	8.34641	9.24532	11.22351
25	25	7.05272	7.78018	8.21457	9.03053	10.10404	11.40867	12.83660	13.82748	15.99525
50	50	8.21457	8.90780	9.32261	10.10403	11.13707	12.40017	13.79114	14.76083	16.89148
100	100	9.32261	9.98630	10.38453	11.13707	12.13657	13.36516	14.72525	15.67679	17.77569
200	200	10.38453	11.02350	11.40795	12.13657	13.10824	14.30798	15.64197	16.57810	18.65034
500	500	11.73036	12.34301	12.71279	13.41591	14.35168	15.52608	16.83093	17.74947	19.78918
1000	1000	12.71279	13.30944	13.67028	14.35768	15.28085	16.42973	17.71652	18.62317	20.64808

For given N, k, (1 ≤ k ≤ N) and α the entries in the above table are the values of y for which $\frac{N!}{(k-1)!(N-k)!} \int_0^y G_r^{k-1}(z) \left[1-G_r(z)\right]^{N-k} g_r(z)dz = \alpha$, where $g_r(z)$ and $G_r(z)$ refer to the p.d.f. and c.d.f. of the standardized gamma chance variable, respectively.

c. The Percentage Points of the Order Statistics

If $B_\alpha(k, N-k+1)$ denotes the α-percentage point of the beta distribution, then y_α , the α-percentage point of the kth order statistic is the solution of

$$G_r(y_\alpha) = B_\alpha(k, N-k+1). \tag{4.16}$$

Equation (4.16) can be written as

$$-y_\alpha + \log_e\left(\sum_{j=0}^{r-1}\frac{(y_\alpha)^j}{j!}\right) = \log_e[1 - B_\alpha(k, N-k+1)] \tag{4.17}$$

and this equation was used to find y_α by an iterative procedure. The values of $B_\alpha(k, N-k+1)$ were obtained from [4]; for those values of k and $N-k+1$ for which $B_\alpha(k, N-k+1)$ was not available, inverse interpolation was performed in the tables of the incomplete beta function [5]. The latter is not necessary for $k=1$ and $k=N$ since for these cases we have, respectively,

$$G(y_\alpha) = 1 - (1-\alpha)^{1/N} \tag{4.18}$$

and

$$G(y_\alpha) = (\alpha)^{1/N} \tag{4.19}$$

5. Best Linear Unbiased Estimates

We consider the gamma p.d.f. in the form

$$g_r(x, \theta) = e^{-x/\theta}\frac{x^{r-1}}{\theta^r\Gamma(r)}, \qquad x > 0, \qquad \theta > 0 \tag{5.1}$$

where θ is to be estimated (r assumed to be known). The best linear unbiased estimate of θ based on the ordered observations is well-known for the case of a complete sample and is given by $\Sigma x/r$. However, if we have a censored sample

$$x_{(r_1)} \leqq x_{(r_1+1)} \leqq \cdots \leqq x_{(r_2)}\ , r_2 \geqq r_1 (r_2\ , r_1 \quad \text{both known}) \tag{5.2}$$

then the problem reduces to determining the vector of constants, $b = (b_1\ , b_2\ , \cdots\ , b_{r_2-r_1+1})$ such that

$$b'\overset{*}{\sum_{x(k)}} b = \text{minimum} \tag{5.3}$$

and

$$b'Ex_{(k)} = \theta \tag{5.4}$$

where $\sum^*_{x(k)}$ denotes the covariance matrix of the ordered x's in (5.2) and $x'_{(k)}$ denotes the row vector $(x_{(r_1)}\ , x_{(r_1+1)}\ \cdots\ x_{(r_2)})$. Since $Ex_{(k)} = \theta\ \mu'_1\ (k, N)$, the constants $b_1\ , b_2\ , \cdots\ , b_{r_2-r_1+1}$ are the solution of

$$\begin{cases} \dfrac{\partial f}{\partial b_j} = 0 \qquad j = r_1\ , r_1+1, \cdots\ , r_2 \\ \displaystyle\sum_{k=1}^{r_2-r_1+1} b_k\mu'_1(k, N) = 1 \end{cases} \tag{5.5}$$

where

$$f = b' \sum_{x(k)}^{*} b - \beta\left\{\sum_{k=1}^{r_2-r_1+1} b_k \mu_1'(k, N) - 1\right\}.$$

The constants b's are given by

$$\begin{bmatrix} b_{r_1} \\ b_{r_1+1} \\ b_{r_2} \\ \beta \end{bmatrix} = S^{-1} \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix} \tag{5.6}$$

where S is the matrix of coefficients on the left side of (5.5). For example, if $r_1 = 1, r_2 = 2, N = 3$, then b_1 and b_2 are given by

$$\begin{bmatrix} b_1 \\ b_2 \\ \beta \end{bmatrix} = \begin{bmatrix} c_{11} & c_{12} & -\frac{\mu_1'(1,3)}{2} \\ c_{12} & c_{22} & -\frac{\mu_1'(2,3)}{2} \\ \mu_1'(1,3) & \mu_2'(2,3) & 0 \end{bmatrix}^{-1} \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \tag{5.7}$$

where c_{11} , c_{22} , c_{12} are the variances and covariance of $x_{(1)}$, $x_{(2)}$ in a sample of size 3 from a gamma distribution.

Some constants b's for the cases of $r = 2, 3$ will be published later. For $r = 1$, i.e. for the case of the exponential distribution these have been obtained in [8].

6. Applications to Life-Testing, Reliability and Other Problems and Some Further Remarks

a. Time to rth Failure

In the following we shall assume the exponential density (5.1) and denote by $x_{(r)}$ the time to rth failure in a life-test experiment with replacements and by n the number of units on test. Then $x = nx_{(r)}/\theta$ has the p.d.f. $g_r(x)$ of (1.1). Thus, the results of this paper will be useful in finding the probability that the smallest (or any other ordered value) time to rth failure in N life-test experiments (with replacements) with n units in each will at least be as large as a given value. Also, the expected values and other moments of the ordered times to rth failure can be found with the help of Table I.

Illustration 1

Life-tests (with replacements) are to be carried on with $n = 10$ units each of the same type, in $N = 3$ test sets. We would like to know the probability that in none of the 3 sets four failures will take place in less than $c = \frac{1}{2}$ times the average life-time. The required probability is $P = 1 - F_4(2; 1, 3)$ which we can find from a table of the Poisson distribution (the parameter of the Poisson distribution being $nc = 2$) as indicated in formula (4.13). Thus, $P = (.857123)^3 =$

TABLE III D

r = 4

This table gives the α-percentage points of the k^{th} smallest order statistic of a sample of size N [1 ≤ k ≤ N, N = 1(1)10, 25, 50, 100, 200, 500, 1000, and selected values of α] from a standardized gamma distribution with parameter r.

N	k \ α	.01	.05	.10	.25	.50	.75	.90	.95	.99
1	1	.82325	1.36632	1.74477	2.53532	3.67206	5.10943	6.68078	7.75366	10.04512
2	1	.67269	1.09427	1.37838	1.95001	2.73162	3.67206	4.65871	5.31458	6.68078
	2	1.74477	2.41115	2.83667	3.67206	4.80885	6.20322	7.71463	8.74903	10.97414
3	1	.59931	.96562	1.20850	1.68825	2.32766	3.07712	3.84580	4.34905	5.38194
	2	1.44477	1.95784	2.27604	2.88199	3.67206	4.59606	5.55268	6.18615	7.50625
	3	2.37197	3.06979	3.50367	4.33981	5.45977	6.82355	8.30184	9.31638	11.50771
4	1	.55271	.88522	1.10343	1.52947	2.08803	2.73162	3.38180	3.80301	4.65871
	2	1.28893	1.72923	1.99825	2.50279	3.14730	3.88467	4.63235	5.11988	6.11932
	3	1.98885	2.52881	2.85562	3.46742	4.25854	5.16198	6.09930	6.71950	8.01337
	4	2.83667	3.54418	3.97861	4.80885	5.91341	7.25519	8.71138	9.71291	11.88236
5	1	.51934	.82826	1.02952	1.41919	1.92412	2.49867	3.07265	3.44163	4.23033
	2	1.18735	1.58265	1.82182	2.26592	2.82581	3.45755	4.08988	4.49826	5.32712
	3	1.78078	2.24346	2.51969	3.02978	3.67206	4.40004	5.13458	5.61262	6.59251
	4	2.40379	2.95457	3.28393	3.89534	4.67390	5.57120	6.49434	7.10527	8.38113
	5	3.20283	3.91264	4.34544	5.16869	6.26022	7.58527	9.02519	10.01729	12.17089
6	1	.49373	.78489	.97351	1.33642	1.80244	2.32766	2.80291	3.18007	3.84580
	2	1.11367	1.47750	1.69602	2.09896	2.60208	3.16410	3.72148	4.07895	4.79947
	3	1.64215	2.05621	2.30126	2.74982	3.30792	3.93233	4.55447	4.95542	5.76868
	4	2.16119	2.63366	2.91258	3.42362	4.06222	4.78232	5.50698	5.97820	6.94403
	5	2.73803	3.29338	3.62309	4.23214	5.00411	5.89141	6.80361	7.40736	8.66968
	6	3.50367	4.21279	4.64323	5.45977	6.54039	7.85207	9.27951	10.26406	12.40528
7	1	.47316	.75026	.92898	1.27111	1.70724	2.19497	2.67457	2.97943	3.58715
	2	1.05672	1.39689	1.60007	1.97262	2.43650	2.94617	3.45009	3.77166	4.41615
	3	1.54032	1.92013	2.14342	2.54957	3.05068	3.60620	4.15484	4.50615	5.21354
	4	1.99664	2.41930	2.66679	3.11650	3.67206	4.29058	4.90511	5.30072	6.10272
	5	2.47136	2.94863	3.22850	3.73673	4.37331	5.08667	5.80328	6.26907	7.22414
	6	3.01728	3.57424	3.90343	4.50941	5.27526	6.15414	7.05738	7.65544	8.90700
	7	3.75832	4.46544	4.89340	5.70376	6.77492	8.07569	9.49260	10.47143	12.60245
8	1	.45611	.72168	.89235	1.21766	1.62996	2.08803	2.53583	2.81921	3.36180
	2	1.01070	1.32850	1.52344	1.87246	2.30235	2.77588	3.23945	3.53397	4.12171
	3	1.46097	1.81487	2.02190	2.39666	2.85589	3.36151	3.85757	4.17360	4.80660
	4	1.87438	2.26164	2.48705	2.89420	3.39322	3.94388	4.46626	4.83310	5.61318
	5	2.28791	2.71510	2.96363	3.41312	3.95590	4.57938	5.18770	5.57906	6.37223
	6	2.73315	3.21272	3.49265	4.00133	4.63198	5.33933	6.04940	6.51078	7.45699
	7	3.25662	3.81373	4.14190	4.74474	5.50503	6.37657	7.27251	7.86634	9.10114
	8	3.97861	4.68317	5.10861	5.91341	6.97648	8.26767	9.67627	10.65012	12.77203
9	1	.44163	.69751	.86142	1.17282	1.56546	1.99935	2.42130	2.68739	3.21366
	2	.97155	1.27877	1.46032	1.79030	2.19477	2.63791	3.06961	3.34299	3.88641
	3	1.39688	1.73004	1.92430	2.27442	2.69696	3.16885	3.62498	3.91445	4.49175
	4	1.77838	2.13880	2.34761	2.72302	3.18025	3.68155	4.17215	4.48437	5.10920
	5	2.15033	2.54168	2.76808	3.17522	3.67206	4.21865	4.75593	5.09925	5.78980
	6	2.53546	2.96506	3.21393	3.66258	4.21257	4.82144	5.42452	5.81230	6.59333
	7	2.95948	3.43997	3.71955	4.22836	4.86320	5.55565	6.25968	6.71748	7.65547
	8	3.46580	4.02223	4.34928	4.94892	5.70415	6.56952	7.45897	8.04871	9.28851
	9	4.17248	4.87421	5.29733	6.09695	7.15300	8.43650	9.83751	10.80708	12.92908
10	1	.42908	.67667	.83485	1.13440	1.51050	1.92412	2.32459	2.57639	3.07265
	2	.93247	1.22880	1.40683	1.72123	2.10479	2.52309	2.92893	3.18509	3.69283
	3	1.36412	1.65927	1.86628	2.17392	2.57502	3.01193	3.43648	3.70505	4.23879
	4	1.70143	2.03925	2.24595	2.58551	3.01041	3.47378	3.92494	4.21094	4.78102
	5	2.04149	2.40556	2.61525	2.99068	3.44609	3.94368	4.43014	4.73932	5.35780
	6	2.38592	2.77966	3.00648	3.41312	3.90779	4.45071	4.98371	5.32410	6.00858
	7	2.75072	3.18159	3.43037	3.87780	4.42504	5.02994	5.60800	6.01322	6.78978
	8	3.15863	3.63927	3.91826	4.42308	5.04642	5.74407	6.39080	6.89885	7.76082
	9	3.65137	4.20672	4.53245	5.12910	5.87976	6.73962	7.62401	8.22713	9.47549
	10	4.34544	5.04424	5.46524	6.26022	7.30999	8.58646	9.98097	10.94696	13.05655
25	25	4.54813	5.14540	5.50628	6.19170	7.10651	8.23505	9.48767	10.36591	12.30861
50	50	5.50628	6.08801	6.43922	7.10651	7.99876	9.10309	10.33359	11.19903	13.11912
100	100	6.43922	7.00552	7.34770	7.99876	8.87152	9.95530	11.16718	12.02157	13.92442
200	200	7.34770	7.90014	8.23442	8.87152	9.72774	10.79414	11.99023	12.83528	14.72029
500	500	8.51577	9.05283	9.37843	10.00031	10.83852	11.88592	13.06424	13.89855	15.97376
1000	1000	9.37843	9.90592	10.22614	10.83852	11.66545	12.70105	13.86856	14.69547	16.55353

For given N, k, (1 ≤ k ≤ N) and α the entries in the above table are the values of y for which $\frac{N!}{(k-1)!(N-k)!}\int_0^y G_r^{k-1}(z)\left[1-G_r(z)\right]^{N-k} g_r(z)dz = \alpha$, where $g_r(z)$ and $G_r(z)$ refer to the p.d.f. and c.d.f. of the standardized gamma chance variable, respectively.

TABLE III E

r = 5

This table gives the α-percentage points of the k^{th} smallest order statistic of a sample of size N [1 ≤ k ≤ N, N = 1(1)10, 25, 50, 100, 200, 500, 1000, and selected values of α] from a standardization gamma distribution with parameter r.

N	k \ α	.01	.05	.10	.25	.50	.75	.90	.95	.99
1	1	1.27911	1.97015	2.43259	3.36860	4.67091	6.27443	7.99359	9.15352	11.60463
2	1	1.07859	1.62909	1.98508	2.67891	3.59649	4.67091	5.77549	6.50055	7.99359
	2	2.43259	3.22364	3.71785	4.67091	5.94204	7.47399	9.11148	10.22207	12.59055
3	1	.97876	1.46445	1.77335	2.36426	3.12580	3.99425	4.86685	5.43081	6.57468
	2	2.06703	2.68826	3.06515	3.77010	4.67091	5.70588	6.76231	7.45537	8.88685
	3	3.17778	3.98586	4.48034	5.42049	6.66025	8.14850	9.74280	10.82837	13.15516
4	1	.91450	1.36023	1.64072	2.17098	2.84312	3.59649	4.34196	4.81865	5.77549
	2	1.87399	2.41383	2.73643	3.33069	4.07458	4.91059	5.74622	6.28596	7.38247
	3	2.72523	3.36102	3.73971	4.43921	5.32315	6.33241	7.36062	8.03562	9.43296
	4	3.71785	4.52624	5.01617	5.94204	7.15753	8.61577	10.18176	11.25109	13.55094
5	1	.86804	1.28570	1.54654	2.03550	2.64799	3.32589	3.98913	4.40993	5.29822
	2	1.74676	2.23594	2.52539	3.05324	3.70533	4.42801	5.14100	5.59710	6.51436
	3	2.47601	3.02680	3.35038	3.93999	4.67091	5.48769	6.30219	6.82809	7.89772
	4	3.21503	3.85361	4.23056	4.92259	5.79236	6.78264	7.79100	8.45366	9.82788
	5	4.13804	4.94205	5.42678	6.33981	7.53612	8.97208	10.51742	11.57503	13.85539
6	1	.83210	1.22850	1.47463	1.93307	2.50208	3.12580	3.67890	4.11204	4.86685
	2	1.65370	2.10727	2.37366	2.86607	3.44627	4.09379	4.72671	5.12875	5.93165
	3	2.30835	2.80536	3.09480	3.61755	4.25788	4.96419	5.65964	6.10429	6.99914
	4	2.92973	3.48295	3.80532	4.38949	5.10999	5.91264	6.71211	7.22836	8.27909
	5	3.60391	4.24133	4.61555	5.30023	6.15811	7.13347	8.12688	8.78014	10.13712
	6	4.48034	5.27860	5.75830	6.66025	7.84107	9.25946	10.78882	11.83733	14.10250
7	1	.80304	1.18255	1.41710	1.85175	2.38725	2.96963	3.53042	3.88219	4.57489
	2	1.58129	2.00797	2.25717	2.70589	3.25329	3.84396	4.41954	4.78332	5.50564
	3	2.18425	2.64323	2.90872	3.38519	3.96395	4.59645	5.21375	5.60588	6.38926
	4	2.73451	3.23318	3.52139	4.03935	4.67091	5.35554	6.04863	6.48530	7.36436
	5	3.29402	3.84679	4.16735	4.74617	5.45767	6.24931	7.03703	7.54577	8.58221
	6	3.92565	4.56029	4.93169	5.60952	6.45726	7.42046	8.40183	9.04772	10.39107
	7	4.76827	5.56056	6.03567	6.92802	8.09575	9.49995	11.01634	12.05753	14.31023
8	1	.77880	1.14442	1.36952	1.78484	2.29353	2.84312	3.36920	3.69772	4.34196
	2	1.52242	1.92323	2.16301	2.58616	3.09608	3.64767	4.17983	4.51468	5.17667
	3	2.08096	2.51704	2.76458	3.20669	3.74002	4.31889	4.88009	5.23476	5.99955
	4	2.58846	3.04820	3.31233	3.78416	4.35494	4.97716	5.58374	5.96891	6.82870
	5	3.07911	3.57737	3.86403	4.37756	5.00190	5.68735	6.36077	6.79126	7.65813
	6	3.59827	4.14933	4.46783	5.04168	5.74580	6.52780	7.30614	7.80888	8.83370
	7	4.19942	4.83073	5.19927	5.87096	6.70998	7.66286	8.63449	9.27483	10.59658
	8	5.01617	5.80265	6.27353	7.15753	8.31423	9.70614	11.21207	12.24712	14.48879
9	1	.75811	1.11201	1.32917	1.72846	2.21497	2.73775	3.23551	3.54527	4.15040
	2	1.47210	1.86132	2.08616	2.48747	2.96939	3.48789	3.98565	4.29782	4.91255
	3	2.00796	2.41481	2.64828	3.06324	3.55635	4.09922	4.61769	4.94408	5.58985
	4	2.47312	2.90325	3.14921	3.58694	4.11224	4.68163	5.23314	5.58164	6.27418
	5	2.91689	3.37601	3.63866	4.10650	4.67091	5.28515	5.88338	6.26319	7.02227
	6	3.36876	3.86567	4.15071	4.66020	5.27835	5.95599	6.62151	7.04697	7.90404
	7	3.85926	4.40805	4.72454	5.29378	5.99117	6.76556	7.53554	8.03342	9.04775
	8	4.43738	5.06513	5.43106	6.09710	6.92844	7.87273	8.83584	9.47096	10.79864
	9	5.23350	6.01443	6.48157	7.35805	8.50529	9.88726	11.38376	12.41354	14.65471
10	1	.74013	1.08395	1.29435	1.67996	2.14775	2.65799	3.12219	3.41640	3.98913
	2	1.42162	1.79881	2.02025	2.40415	2.86298	3.35435	3.82413	4.11778	4.69437
	3	1.96742	2.32913	2.57875	2.94477	3.41481	3.91950	4.40409	4.70816	5.30767
	4	2.38022	2.78521	3.02974	3.42700	3.91776	4.44643	4.95587	5.27654	5.91119
	5	2.78787	3.21710	3.46157	3.89511	4.41500	4.97716	5.52123	5.86495	6.54813
	6	3.19411	3.65204	3.91325	4.37756	4.93659	5.54415	6.13557	6.51104	7.26155
	7	3.61859	4.11378	4.39715	4.90286	5.51555	6.18666	6.82302	7.20662	8.11187
	8	4.08754	4.63385	4.94836	5.51336	6.20487	6.97219	7.67835	8.23014	9.16123
	9	4.64752	5.27181	5.63515	6.29613	7.12073	8.05745	9.01353	9.66262	10.99810
	10	5.42678	6.20246	6.66826	7.53612	8.67498	10.04798	11.53640	12.56177	14.78816
25	25	5.81761	6.48439	6.88449	7.63950	8.63879	9.86691	11.20653	12.14442	14.20609
50	50	6.88449	7.52566	7.91075	8.63879	9.60588	10.79446	12.10998	13.03056	15.06182
100	100	7.91075	8.52889	8.90088	9.60588	10.54591	11.70644	12.99674	13.90246	15.90776
200	200	8.90088	9.49931	9.86024	10.54591	11.46333	12.60031	13.86928	14.76240	16.74606
500	500	10.16339	10.74055	11.08957	11.75449	12.64750	13.75886	15.00395	15.88266	17.83963
1000	1000	11.08957	11.65370	11.99543	12.64750	13.52530	14.62071	15.85111	16.72001	18.66580

For given N, k, (1 ≤ k ≤ N) and α the entries in the above table are the values of y for which $\frac{N!}{(k-1)!(N-k)!}\int_0^y G_r^{k-1}(z)\left[1-G_r(z)\right]^{N-k} g_r(z)dz = \alpha$, where $g_r(z)$ and $G_r(z)$ refer to the p.d.f. and c.d.f. of the standardized gamma chance variable, respectively

.6297. Also from Table III, we can put one-sided or two-sided confidence limits on the smallest (or any other) time to rth failure in terms of the average life or vice versa. For the present example $r = 4$ and from Table IIID we find that the 90% lower confidence limit on $x_{(1)}$, the smallest time to the fourth failure is .1209 times the average life θ. Again the 90% lower confidence limit on θ based on $x_{(1)}$, say, is $(10/3.8458)\, x_{(1)} = 2.6002\, x_{(1)}$.

Again by using Table I we find,

Expected value of the smallest time to the 4th failure = .2449 av. life

Expected value of the median time to the 4th failure = .3820 av. life

Expected value of the largest time to the 4th failure = .5731 av. life

If we have N systems and the rth failure of a unit in any of these means that the system becomes unsatisfactory or unreliable (replacements are made till the $(r-1)$th failure), then the probability that 0, 1, 2, $\cdots$ N of the systems arrive at an unsatisfactory state before a certain time can be obtained from the formulae of Section 4 and also individual and joint confidence limits can be placed on the reliability of these systems.

b. Total Life

If $t_1, t_2, \cdots, t_r$ are the times to 1st, 2nd, $\cdots$, rth failures in n units on test (without replacement) that follow the exponential law $f(t) = \lambda e^{-\lambda t}$, then the total life T defined by

$$T = t_1 + t_2 + \cdots + t_r + (n-r)t_r \tag{6.1}$$

has the gamma density function

$$g_r(T) = \lambda^r e^{\lambda T} \frac{T^{r-1}}{\Gamma(r)}, \qquad T > 0 \tag{6.2}$$

so that results of this paper will be applicable in comparing two or more populations using total life as the criterion.

The distributions of the statistic T when t's come from exponential and uniform distribution are derived in [1] and [2], respectively.

Illustration 2

The most probable value of the largest order statistic is of importance in the study of breakdown voltages of paper capacitors of several layers. It is assumed that the size of the conducting particles (impurities) follows an exponential distribution and that the particles of one layer are lined up against those of the next layer. Under this assumption the total length of the r conducting particles follows a gamma distribution. We are then interested in the mode of the largest of a sample of N from a gamma distribution.

Let $N = 10$ be the number of particles per square unit area on each of the $r = 3$ layers. Then from Table II, we find that the modal value of the largest length of the conducting particle is equal to $5.448/\lambda$ where λ is the parameter of the exponential distribution $\lambda e^{-\lambda x}$.

c. Relation to Order Statistics from χ^2-Distribution

The random variable $\frac{1}{2}\chi^2_\nu$, with 2ν degrees of freedom has the gamma density function (1.1) with parameters $r = \nu$. If y and y' denote the kth order statistic of a sample of size N from the standarized gamma distribution with parameter r and χ^2_{2r} (with $2r$ degrees of freedom) respectively, then we have

$$P[y' \leqq c] = P\left[y \leqq \frac{c}{2}\right] \tag{6.3}$$

$$y'_\alpha = 2y_\alpha \tag{6.4}$$

where y'_α and y_α refer to the α-percentage points of y' and y respectively. Also

$$E(y')^i = 2^i E y^i. \tag{6.4}$$

d. Application to Distribution of Ordered Sample Variances from Normal Populations

If s^2_ν is the sample variance based on ν degrees of freedom from a normal distribution with variance σ^2, then $\nu s^2_\nu/\sigma^2$ is distributed like a χ^2_ν with ν degrees of freedom. This enables us to obtain the moments and distribution of the ordered sample variances for $\nu = 2r$ degrees of freedom from the results and tables of the present paper.

7. Acknowledgments

The author is thankful to Mr. G. J. Levenbach for bringing a problem to his attention which led to the writing of the present paper. The author also wishes to thank Mr. R. B. Murphy for a constructive review and comments and Messrs. J. Nadler, M. Sobel, M. E. Terry, J. A. Tischendorf, and J. W. Tukey for commenting on an earlier version of this paper. Miss Phyllis A. Groll programmed and carried out the computations of the tables in the present paper and her help is thankfully acknowledged.

References

[1] Epstein, B. and Sobel, M., "Some Theorems Revelant to Life Testing from an Exponential Distribution", *Annals of Mathematical Statistics*, Vol. 25 (1954), pp. 373–381.

[2] Gupta, S. S. and Sobel, M., "On the Distribution of a Statistic Based on Ordered Uniform Chance Variables", *Annals of Mathematical Statistics*, Vol. 29 (1958), pp. 274–281.

[3] Molina, E. C., *'Poisson's Exponential Binomial Limit'* Van Nostrand Company, New York, 1942.

[4] Pearson, E. S. and Hartley, H. O., *Biometrika Tables for Statisticians*, Vol. I, University Press, Cambridge (England), 1954.

[5] Pearson, K., *Tables of the Incomplete Beta-Function*, University Press, Cambridge (England), 1948.

[6] Pearson, K., *Tables of the Incomplete Γ-Function*, University Press, Cambridge (England), 1957.

[7] Sarhan, A. E., "Estimation of the Mean and Standard Deviation by Order Statistics, Part III", *Annals of Mathematical Statistics*, Vol. 26 (1955), pp. 576–592.

Parallel Fractional Replicates

CUTHBERT DANIEL

New York City

In studying the effects of many factors on several responses the experimenter is sometimes able to say which factors are likely to be influential for each response. This knowledge permits more efficient planning since certain effects and interactions are now negligible. Some 2^{p-q} fractional replicates are constructed that take advantage of this knowledge.

Most industrial products have more than one important property. Experiments to measure the influence of process variables on several properties are often carried on simultaneously. This is likely to be the case when a large part of the cost of experimentation is spent in producing material for tests. Thus even small batches of a metal, of a plastic, or of a new food product may be more expensive to prepare than to test for several properties. As a simple example, imagine that three factors affecting the flavor (y_1) and texture (y_2) of a new food product are being studied. An essential oil (A) is to be added that may modify the flavor (y_1) although it cannot influence the texture (y_2). A wheat product (B) that may influence the flavor *and* texture is also varied. Finally, a bran-like material (C) without flavor but likely to influence the texture is to be the third factor. This state of affairs may be abbreviated as "influence pattern" A–B–C. or as 1-1-1. The first term in such an abbreviation indicates the number of factors that are taken to influence only y_1. The second term numbers the factors that influence both y_1 and y_2. The third term shows the factors that influence y_2 only.

The familiar half-replicate, 2^{3-1}, with treatment combinations (1), ab, ac and bc can be viewed as a full 2^2 factorial experiment on A and B for studing their effects and interaction on y_1. At the same time this plan is a full 2^2 on B and C and so will give information on their effects and interaction on y_2. Six estimates are obtained then from four "runs". If the assumptions made about the influence pattern are correct, then effects due to C, AC, and BC do not exist for y_1, and effects due to A, AB, and AC are zero for y_2. There is then no confounding in this fractional replicate.

Proceeding to slightly more realistic examples, assume that the three factors A, B and C are known to influence y_1, and that three other factors D, E and F influence y_2. Assume further that each set of factors influences only one response. The influence pattern is now written 3–0–3. The obvious plan would be to adjoin a 2^3 on A, B and C to one on D, E and F. The combined set of treatments might look like this:

(1)
a	*d*
b	*e*
ab	*de*
c	*f*
ac	*df*
bc	*ef*
abc	*def*

The alias subgroup of this 2^{6-3} includes AD, BE and CF. If the assumptions that A, B and C do not influence y_2 and that D, E and F do not influence y_1 are all correct, there is no confounding here. But the results will not give any clue to the failure of one or more of these assumptions. The two factorials can be adjoined in many ways. It will be better to look for a combination that is less sensitive to these assumptions.

Since for eight runs the combined experiment must be a 2^{6-3}, there will be seven interactions in the alias subgroup and there will be twenty-four letters. The average length of the interactions is therefore 3 3/7 letters, and so some interactions must be three-letter ones. There must then be some confounding of main effects with two-factor interactions. The alias subgroup

$$\begin{array}{lll} -A\underline{DE} & AB\underline{DF} & -AC\underline{F} \\ -B\underline{EF} & ABC\underline{E} & I \\ -BC\underline{D} & C\underline{DEF} & \end{array}$$

guarantees that main effects on y_1 are confounded only with two-factor interactions on y_2, or on y_1 and y_2 together. (The letters D, E, F are underlined.) Also estimates of main effects of factors D, E and F on y_2 are only confounded with two-factor interactions on y_1, or with mixed two-factor interactions.

A simpler picture of what confounding is taking place is given by listing the sums of effects and 2 f.i. (two-factor interactions) that will be estimated by each of the familiar orthogonal contrasts:

$$\begin{array}{lll} A - DE - CF & D - AE - BC & AB + CE + DF \\ B - EF - CD & E - AD - BF & \\ C - BD - AF & F - BE - AC & \end{array}$$

The seven contrasts in the y_1 measurements will be given the interpretations A, B, C, BC, error, AC and AB respectively. The same contrasts in the y_2 measurements are given the interpretations DE, EF, error, D, E, F and DF respectively. The eight corresponding treatment combinations or experimental conditions are (1), $\underline{aef}$, $\underline{bde}$, $abdf$, $\underline{acd}$, $cdef$, $abce$, bcf, generated here by the three underlined combinations.

The influence pattern 2-1-2 (meaning that two factors influence y_1 only, that one factor influences both responses, and that two factors influence y_2

only) might also be studied in eight runs in a 2^{5-2} with alias subgroup $-ACE$, $-BCD$, $ABDE$, I. The confounding pattern is given more explicitly by:

	y_1	y_2
$A - CE$	A	$-CE$
$B - CD$	B	$-CD$
$C - AE^* - BD^*$	C	C
$D - BC$	$-BC$	D
$E - AC$	$-AC$	E
$AB + DE$	AB	DE
$AD^* + BE^*$	error	error.

The starred interactions are assumed to be zero. The eight experimental conditions are (1), $\underline{abc}$, $\underline{ae}$, bce, $\underline{bd}$, acd, $abde$, cde, The seven contrasts among the y_1 measurements will be given the interpretation shown in the second column above; the same contrasts in the y_2 measurements will be given the names shown in the third column.

For influence pattern 1-3-1 a plan in eight runs is possible although the 2 f.i. will be confounded in pairs for each response. The alias subgroup I, $ABCD$, $BCDE$, AE produces the experimental conditions (1), $\underline{abe}$, $\underline{bc}$, ace, $\underline{cd}$, $abcde$, bd, ade. The seven sums of confounded effects are now:

$A + E$	$AB + CD + BE$
$B +$ error	$AC + BD + CE$
$C +$ error	$AD + BC + DE$.
$D +$ error	

Although these plans are still too small for everyday use, Table 1 shows three two-factor-interaction-clear (2 f.i.c.) plans and four that confound 2 f.i. in pairs. These plans are not robust against wrongly assumed influence patterns and they give little or no information about error.

Table 2 summarizes some 16-run, symmetrical influence pattern, bi-variate, 2 f.i.c. plans. The column headed 3-1-3 is expanded here to show in more detail what confounding occurs and what sets of experimental conditions are required. It is known in this case that three factors, A, B and C affect only y_1 , that one factor, D, affects both y_1 and y_2 , and that three other factors, E, F and G affect only y_2 . Multiplying out all products of the three generators of the alias subgroup gives

1.	$ABCG$	$1. \times 2. CDFG$	$1. \times 2. \times 3. AEFG$
2.	$ABDF$	$1. \times 3. BDEG$	
3.	$ACDE$	$2. \times 3. BCEF$	

TABLE 1

Eight-run, symmetric-influence-pattern, bivariate, 2^{p-q} plans
I. Two-factor interaction clear
II. Main effects clear, 2 f.i. confounded in pairs

I. Influence Patterns	1-2-1	2-1-2	3-0-3	
Factor Letters	*A-BC-D*	*AB-C-DE*	*ABC--DEF*	
p	4	5	6	
q	1	2	3	
Members of Alias Subgroups	$\underline{ABCD}^*$	$-\underline{ACE}$ $-\underline{BCD}$ $(\underline{ABDE})$	$-\underline{ADE}$ $-\underline{BEF}$ $-\underline{BCD}$	$(ABDF)$ $(ABCE)$ $(CDEF)$ $(-ACF)$
Generators of Principal Blocks	*ab* *ac* *ad*	*abc* *abde* *acd*	*acd* *aef* *bcf*	

II. Influence Patterns	1-3-1	2-2-2	3-1-3	4-0-4	
Factor Letters	*A-BCD-E*	*AB-CD-EF*	*ABC-D-EFG*	*ABCD--EFGH*	
p	5	6	7	8	
q	2	3	4	5	
Members of Alias Subgroups	$\underline{ABCD}^*$ $\underline{BCDE}$ $\underline{AE}$	$-\underline{ACE}$ $-\underline{ADE}$ $-\underline{BCF}$ $(\underline{ABCD})$ $(\underline{CDEF})$	$\underline{AE}$ $\underline{BF}$ $\underline{CG}$ $\underline{ABCD}$ $(\underline{DEFG})$	$\underline{AE}$ $\underline{BF}$ $\underline{CG}$ $\underline{DH}$ $\underline{ABGH}$	$(ABCD)$ $(EFGH)$
Generators of Principal Blocks	*abe* *ace* *ade*	*abcd* *abef* *acf*	*ade* *bdf* *cdg*	*abef* *aceg* *adeh*	

* Generators of alias subgroups underlined; other important members in parentheses.

Thus, since every member of the alias subgroup has four letters, each main effect is confounded only with three-factor interactions. There will be three degrees of freedom for error estimation in each set of response contrasts, namely, those measuring D, E and F in y_1 and those measuring A, B and C in y_2. The twenty-one 2 f.i. are confounded in the seven strings:

$$AB + CG^* + DF \qquad AF^* + BD + EG$$
$$AC + BG^* + DE \qquad AG^* + BC + EF$$
$$AD + BF^* + CE^* \qquad BE^* + CF^* + DG.$$
$$AE^* + CD + FG$$

TABLE 2

Sixteen-run, symmetric-influence-pattern, two-factor-interaction-clear, bivariate, 2^{p-q} plans

Influence Patterns	1-3-1	2-2-2	3-1-3	4-0-4
Factor Letters	*A-BCD-E*	*AB-CD-EF*	*ABC-D-EFG*	*ABCD–EFGH*
p	5	6	7	8
q	1	2	3	4
Generators of Alias Subgroups	$-ABCDE$	*ABCF* *ADEF*	*ABCG* *ABDF* *ACDE*	*ABCE* *ABDF* *ACDG* *BCDH*
Generators of Principal Blocks	*ab* *ac* *ad* *ae*	*abcf* *ace* *adef* *af*	*abcg* *abdf* *acde* *adg*	*abce* *abdf* *acdg* *bcdh*

The nine starred interactions are assumed negligible since their two factors taken together are stated not to influence y_1 or y_2.

The set of 16 experimental conditions are produced from the four "generators of the principal block" by writing down all their products:

1. *abcg*	1. × 2. *cdfg*	2. × 4. *bfg*	1. × 3. × 4. *abe*	
2. *abdf*	1. × 3. *bdeg*	3. × 4. *ceg*	2. × 3. × 4. *abcdefg*	
3. *acde*	1. × 4. *bcd*	1. × 2. × 3. *aefg*	1. × 2. × 3. × 4. *def*	
4. *adg*	2. × 3. *bcef*	1. × 2. × 4. *acf*		(1).

These combinations are easier to see when written in alphabetical order, thus:

abcdefg	*bcd*
abcg	*bcef*
abdf	*bdeg*
abe	*bfg*
acde	*cdfg*
acf	*ceg*
adg	*def*
aefg	(1).

This plan is still not very robust against a wrongly assumed influence pattern. There are only four "error" degrees of freedom for each property, whose irregularity might indicate such a failure. However the improvement over the eight-run plans is noticeable. Clearly 32-run plans will give more "error" degrees of freedom, some of which may show significant inflation in case of serious mistakes in the assumed influence pattern.

Unsymmetrical influence patterns are common enough to be worth study. If the pattern is very skew, e.g. 4-1-1, then the requirements for y_1, the response with the largest number of factors, dominate the plan. But moderate asymmetry can sometimes be adjusted nicely to the corresponding asymmetry in confounding of many 2^{p-q} plans.

Extensions to three or more properties do not seem easy to classify. They must be worked out in detail for each case. The simplest example is our old friend the 2^{3-1}, used now under the (unlikely) set of assumptions that only A and B affect y_1, that only B and C affect y_2, and that only A and C affect y_3. The 2^{3-1} may be viewed as three complete 2^2 factorials, one in A and B, one in B and C, one in A and C.

A convenient way is needed to summarize the experimenter's knowledge and feelings about the effects of each of K factors on each of R responses. A $K \times R$ "influence matrix" has been useful both in aiding the statistician to understand the limitations and advantages of the experimenter's technical background, and to record the experimenter's state of belief before the new round of experiments is started. Entries of -1, 0, $+1$ can be used to indicate the experimenter's opinions about the sign and magnitudes of real effects. An i can be used to indicate ignorance.

Useful proposals for the analysis of multivariate experiments seem to me to wait upon a more explicit classification of the types of responses to be measured in each run. Thus some responses are of similar properties, measured in the same units. Tensile strength before and after aging might be an example of such bi-variates. The second type of paired responses are similar or at least related physical or chemical properties, not of the same dimensions. Flexural strength and flexural modulus are a pair of this kind. Third might come properties that do not depend on the same basic attributes of the material being produced. Brittleness and electrical conductivity, flavor and color, corrosion resistance and machinability are pairs that come in this category. We must expect different sorts of dependence, functional and probabilistic, for these classes of responses. A single statistical model will hardly apply to all.

Detailed numerical examples will be published later.

NOTICES

DEPARTMENT OF STATISTICS
THE FLORIDA STATE UNIVERSITY

To assist in meeting the demands for well trained statisticians, the Florida State University in 1959 founded a new Department of Statistics with immediate authorization to offer programs leading to the Bachelor of Science and Master of Science degrees in statistics. With the further development of the department, programs leading to the Doctor of Philosophy degree are anticipated.

The Bachelor of Science degree in statistics is designed to prepare the student for continuing graduate work, for employment related to statistical computations, and for employment with statistics groups in government, industry and business. A major in statistics consists of a minimum of eighteen semester hours in statistics courses at the junior and senior college level. Competency in mathematics is essential and a minimum of eighteen semester hours in mathematics is required. Survey courses in areas of application of statistics are recommended and at least twelve semester hours are to be completed in such an area.

The Department of Statistics offers the Master of Science degree in statistics with a thesis requirement. Prior work in statistics is not a requirement for admission to graduate work, but mathematical training comparable to an undergraduate major in mathematics is necessary. Graduates in the sciences may be admitted to graduate work in statistics, but it may be necessary to make up deficiencies in mathematics. The program for the Master of Science degree has been developed so that students completing the program may continue directly in graduate training in statistics toward the Doctor's degree at other universities. The Department of Statistics at the Florida State University offers unusual individual attention and cooperation with the beginning graduate student.

In addition to degree programs, the Department of Statistics is engaged in teaching and consultation in statistics with many other university departments. Courses are offered for other curricula. Faculty and students consult on research projects and theses. Members of the Department of Statistics are engaged in research in statistics in a variety of areas of interest. The editorial office of BIOMETRICS, the journal of the International Biometric Society is in the Department of Statistics.

For complete details on degree requirements, the catalog bulletins of the University should be consulted.

ASSISTANTSHIPS. Graduate fellowships are available at stipends for the academic year of $1500 for a Master's Degree candidate. These awards require no service and provide a full academic year of residence. Fellows pay the regular

resident registration fee, but out-or-state tuition is waived. Applications for fellowships should be filed by February 15th for the following academic year. Later applications will be considered if funds are still available. Fellowships are tax exempt.

Graduate assistantships provide stipends for the academic year varying from $160 to $240 per month for half-time service. Both teaching and research assistantships are available in the Department of Statistics and in association with the Computing Center and other departments. The amount of the stipend is determined by the qualifications of the candidate and the nature and the amount of the part-time services to be rendered. Assistants pay the regular resident registration fee, but out-of-state tuition is waived. Applications for assistantships are considered at the time received, but should normally be filed by March 15 for the following academic year.

Additional information on fellowships and assistantships may be obtained from the Department of Statistics and from the Graduate School. Applications for fellowships are made to the Graduate School and for assistantships to the Department of Statistics.

COURSES. Courses in statistics offered by the Department of Statistics are as follows:

For Undergraduates:

216 Elementary Statistics
316 General Theory of Statistics

For Advanced Undergraduates and Graduates:

406 Statistical Procedures for Research
407 Statistical Methods and Design
426 Fundamentals of Sampling
431 Mathematical Statistics I
432 Mathematical Statistics II
446 Theory of Probability
491-2-3 Directed Individual Study

For Graduates:

501 Theory of Statistics I
502 Theory of Statistics II
506 Advanced Topics in Statistical Methods
507 Nonparametric and Rank Order Statistics
508 Introduction to Multivariate Analysis
509 Advanced Statistical Inference
591-A*B*C Directed Individual Study
599 Master's Thesis

Descriptions and prerequisites for courses with other details are given in the Florida State University general catalogue and Graduate Bulletin.

Additional information and assistance with applications may be obtained

by writing to Dr. Ralph A. Bradley, Head, Department of Statistics, Florida State University, Tallahassee, Florida. Telephone, Tallahassee 2-8000, Extension 2369.

Cornell University Industrial Engineering Seminars

The Department of Industrial and Engineering Administration, Sibley School of Mechanical Engineering, Cornell University, announces the "Cornell University Industrial Engineering Seminars" to be held at Cornell University, Ithaca, New York, June 14 through June 17, 1960. Participants may enroll in one of the following seven groups and also attend general sessions.

Industrial Management
Engineering Administration
Operations Management of the Smaller Company
Work Measurement
Systems Simulation Using Digital Computers
Statistical Decision-Making: Theory and Applications
Statistical Reliability Analysis: Theory and Applications

Specialists from both industry and the staff of Cornell University will provide the seminars. The session is planned for operating management personnel in line supervision and staff positions in industrial engineering, production engineering, engineering administration, research and development, quality control, production control, cost control, materials control, purchasing, and data processing.

For additional information write J. W. Gavett, Seminars Coordinator, Department of Industrial and Engineering Administration, Upson Hall, Cornell University, Ithaca, New York.

Statistical Quality Control for the Textile Industry

The School of Textiles of the North Carolina State College, Raleigh, North Carolina, in cooperation with the College Extension Division will sponsor this short course May 2–13, 1960.

Course Content

First Week, May 2–6, 1960

Monday

Registration and Introduction
Populations and Samples
Frequency Distributions
Methods of Calculation and Interpretation of Standard Deviation
Practice Problems

Tuesday

Calculations of Coefficient of Variation and Applications to Textile Processes
Statistical Quality Control Charts ($\bar{X}$ and R)
Practice Problems and Case Studies

Wednesday

Statistical Quality Control Charts ($\bar{P}$, PN and $\bar{c}$)
Applications, Case Studies and Practice Problems

Thursday

Quality Control Systems by Areas for Textile Plants
Interpretation and Application of Testing Data to Manufacturing Processes

Friday

Applications of Quality Control Techniques with Case Studies
Review with Question and Answer Period

Second Week, May 9–13, 1960

Monday

Tests for Significance of Differences and Confidence Intervals Applied to Textiles

Tuesday

Continuation of the Application of Significance Testing and Confidence Intervals to Textiles, with Practical Problems
Correlation and other Relationships of Variables as Related to Textiles

Wednesday

Applications of Correlation to Textile Problems
Analysis of Variance

Thursday

Case Studies in the Use of Analysis of Variance
Sampling Theory as Applied to Continuous Production and Lot-by-Lot Inspection
Continuation of Sampling Theory and Special Applications of Statistical Techniques to Textile Quality Control Problems

The course will begin on Monday, May 2 and end on Friday, May 13. There will be seven hours of instruction per day which will include lectures, case studies, and practical problems.

In order to provide personal instruction, enrollment is limited to twenty. Registration for the course will be accepted in the order of paid registrations received.

Each participant for the full course will be provided with a notebook containing the necessary materials, as well as a copy of the appropriate textbook. Each enrollee is encouraged to have a slide rule for use in the course.

The instruction for the course will be under the direction of Prof. D. S. Hamby of the Department of Textile Technology, School of Textiles, N. C. State College. Assisting as full-time instructors are Dr. R. J. Hader, Department of Experimental Statistics, N. C. State College, and Prof. W. C. Stuckey, Jr., Department of Textile Technology, School of Textiles, N. C. State College. In addition, guest lecturers will be invited for discussions during the case studies. All of the instructors have had experience in the teaching of statistical quality control, both to students and personnel in short courses. They are also well qualified in industrial applications because of their work as consultants to numerous firms in the textile industry.

Registration will be conducted in the Nelson Building (School of Textiles) from 9:00 to 9:30 A.M. on Monday, May 2, 1960. The registration fee is $160.00 per person for the full course. This fee encludes tuition, supplies, and text materials, but does not include housing accommodations or meals.

The Statistical Summer Seminar

The Statistical Summer Seminar will hold its tenth meeting July 18–25, 1960 at the NUY Gould House, Ardsley-on-Hudson, New York to discuss "The Use of Computers in the Sixties—Boom or Bust?" Dr. Meyer A. Efroymson of Esso Research and Prof. Fred C. Leone of Case Institute of Technology are organizing a program which will include several noted speakers.

This seminar, a private non-profit organization, was organized in 1948 by Profs. Frederick Mosteller and John Tukey to explore new fields of application for statistics. At present it is under the direction of Dr. Milton E. Terry of the Bell Telephone Laboratories.

The registration fee of $175.00 includes comfortable living accommodations. Grants in aid are available to government and academic personnel. For further information write to Dr. Ramanathan Gnanadesikan, Bell Telephone Laboratories, Murray Hill, New Jersey.

Pulp and Paper Statistics Course Announced

The ninth Pulp and Paper Statistics Course will be held August 8th through 19th, 1960, at Brevard College, Brevard, N. C., and will be jointly sponsored by the Institute of Statistics of North Carolina State College, The Technical Section of the Canadian Pulp and Paper Association, and the Technical Association of the Pulp and Paper Industry (TAPPI).

Two courses of instruction, one basic and one advanced, will be held concurrently during the two-week period. The courses will be concerned with the techniques of statistical analysis in research, control and management methods in the pulp and paper industry.

The *Elementary Course* is intended for those people who wish to acquire some knowledge of, and facility with, the simpler but exceedingly valuable statistical techniques and who have had little or no previous formal training

in statistical methods. The training received from this course should prove profitable to organizations now making little or no use of statistical methods but wishing to know more about how they work and how to use them. It will also be profitable to organizations with statistical programs already operating who wish to strengthen their position by having additional men with statistical training available within their organization.

The *Advanced Seminar* is intended for people who have experience and definite facility in the use of statistical techniques and who wish to gain training and practice in special problems that arise in application. The course will consist of laboratory sessions devoted to the consideration of specific problems in the industry, and the practical analysis of these problems.

Further information and application forms for the course may be obtained from the Institute of Statistics, North Carolina State College, Raleigh, N. C.

THE ANNALS OF MATHEMATICS STATISTICS

Vol. 31, No. 1—March, 1960

Contents

The Algebra of a Linear Hypothesis ..... *Henry B. Mann*

An Operational Approach to the r-Way Crossed Classification .. *J. D. Bankier*

Second Order Rotatable Designs in Four or More Dimensions *Norman R. Draper*

A Matrix Substitution Method of Constructing Partially Balanced Designs *B. V. Shah*

Some Approximations to the Binomial Distribution Function .. *R. R. Bahadur*

On the Mixture of Distributions ..... *Henry Teicher*

On Interchanging Limits and Integrals ..... *John W. Pratt*

Moments of the Absolute Difference and Absolute Deviation of Discrete Distributions ..... *S. K. Katti*

Priority Queues ..... *Rupert G. Miller, Jr.*

On Time Dependent Queuing Processes ..... *J. Keilson and A. Kooharian*

A Geometry of Binary Sequences Associated with Group Alphabets in Information Theory ..... *R. C. Bose and Roy R. Kuebler, Jr.*

A Necessary and Sufficient Condition for the Existence of Consistent Estimates ..... *Lucien LeCam and Lorraine Schwartz*

The Distribution of the Latent Roots of the Covariance Matrix . *Alan T. James*

Two-Sample Tests for Multivariate Distributions ..... *Lionel Weiss*

A Modification of the Sequential Probability Ratio Test Analysis to Reduce the Sample Size ..... *T. W. Anderson*

Sequential Tolerance Regions ..... *Sam C. Saunders*

A Useful Generalization of the Stein Two-Sample Procedure . *R. Wormleighton*

Notes:

Sums of Small Powers of Independent Random Variables .. *J. M. Shapiro*

On the Median of the Distribution of Exceedances ..... *K. Sarkadi*

The purpose of the Institute of Mathematical Statistics is to encourage the development, dissemination and application of mathematical statistics. Membership dues, which include a subscription to the *Annals of Mathematical Statistics* are $10.00 per year for residents of the United States of America and $5.00 a year for residents of other countries. Inquiries regarding membership to the Institute should be sent to the Secretary: G. E. Nicholson Jr., Department of Statistics, University of North Carolina, Chapel Hill, North Carolina, USA.

BIOMETRIKA

VOL. 47 PARTS 1 & 2 JUNE 1960

CONTENTS

Memoirs:

BARTLETT, M. S., GOWER, J. C. & LESLIE, P. H. A comparison of theoretical and empirical results for some stochastic population models

KENDALL, DAVID G. Birth-and-death processes, and the theory of carcinogenesis

ASHFORD, J. R., SMITH, C. S. & BROWN, SUSANNAH. The quantal response analysis of a series of biological assays on the same subject

WALKER, A. M. Some consequences of superimposed error in time series analysis

FINCH, P. D. Deterministic customer impatience in the queueing system GI/M/1

BARTON, D. E., DAVID, F. N. & FIX, EVELYN. The polykays of the natural numbers

SRIVASTAVA, A. B. L. The distribution of regression coefficients in samples from bivariate non-normal populations I. Theoretical investigation

KASTENBAUM, MARVIN A. The separation of molecular compounds by countercurrent dialysis: a stochastic process

SAW, J. G. A note on the error after a number of terms of the David-Johnson series for the expected values of normal order statistics

WATSON, G. S. More significance tests on the sphere

JOHNSON, N. L. An approximation to the multinomial distribution: more properties and applications

VAGHOLKAR, M. K. & WETHERILL, G. B. The most economical binomial sequential probability ratio test

MCGREGOR, J. R. An approximate test for serial correlation in polynomial regression

BERKSON, JOSEPH. Nomograms for fitting the logistic function by maximum likelihood

HAIGHT, FRANK A. & BREUER, MELVIN ALLEN. The Borel-Tanner distribution

BURR, E. J. The distribution of Kendall's score S for a pair of tied rankings

Miscellanea:

Contributions by D. H. BHATE, C. R. BLYTH & C. L. CURME, C. BURROWS, C. W. CLUNIES-ROSS & R. H. RIFFENBURGH, P. D. FINCH, R. G. LAHA & E. LUKACS, P. H. LESLIE, H. D. PATTERSON, B. R. RAO, G. L. TURIN, C. A. WILKINS, M. E. WISE.

Reviews:

Other Books received:

The subscription, payable in advance, is 54/- (or $8.00) per volume (including postage). Cheques should be made payable to *Biometrika*, crossed *"A/c Biometrika Trust"* and sent to the Secretary, Biometrika Office, University College, London, W. C. 1. All foreign cheques must be drawn on a Bank having a London agency.

Issued by THE BIOMETRIKA OFFICE, University College London

BIOMETRICS

Journal of the Biometric Society

Vol. 16, No. 2 CONTENTS June 1960

The Role of Statistics in Physiological Research *C. W. Emmens*

Rank Analysis of Incomplete Block Designs: Method of Paired Comparisons Employing Unequal Repetitions on Pairs *Otto Dykstra*

The Treatment of Reciprocal Interaction, with or without Lag, in Path Analysis *Sewall Wright*

Estimating the Parameter in a Conditional Poisson Distribution *A. C. Cohen, Jr.*

Models for the Interpretation of Experiments Using Tracer Compounds *Jerome Cornfield, Jesse Steinfield, and Samuel W. Greenhouse*

Genetic Correlations with Multiple Alleles *R. G. Stanton*

Analysis of Experiments Measuring Threshold Taste *E. K. Harris*

An Unbiased Sampling and Estimation Procedure for Creel Censuses of Fishermen *D. S. Robson*

The Mathematical Theory of Biological Assay of a Local Anaesthetic *N. K. Chakravarti*

QUERIES AND NOTES

Calculation of Inbreeding in Family Selection Studies on the IBM 650 Data Processing Machine *K. Hoen and A. H. E. Grandage*

A Note on Optimum Family Size in Selection Programs *Allan Robertson*

On Heuristic Estimation Methods *Malcolm E. Turner*

A Genetic Application of the Schumann-Bradley Table *Stanley Wearden*

Biometrics is published quarterly by The Biometric Society. Its objects are to describe and exemplify the use of mathematical and statistical methods in biological and related sciences, in a form assimilable by experimenters. The annual non-member subscription rate is $7. Inquiries, orders for back issues and non-member subscriptions should be addressed to:

BIOMETRICS
Department of Statistics
The Florida State University
Tallahassee, Florida

PREPARATION OF MANUSCRIPTS

Manuscripts should be submitted to the office of the editor: J. S. Hunter, Mathematics Research Center, U. S. Army; The University of Wisconsin; Madison, Wisconsin. Each manuscript should be typewritten, double spaced, with wide margins at sides, top, and bottom. The original should be submitted with two additional copies, on paper that will take corrections. Dittoed or mimeographed papers are acceptable only if completely legible. Footnotes should be avoided and replaced by remarks in the text, or placed in an appendix. Preferably, references in the manuscript should appear as (Jones, A. B., 1958), and again later in alphabetical order in a list of references. Alternatively, references may be numbered, e.g. [1], as they appear in the manuscript and be listed in this sequence in the list of references. In the reference list, each reference should contain, in the order indicated, the name and initials of the author followed by those of the co-authors, date of publication, title of reference, source, volume number and page. References to books should include publisher's name and location.

Figures, charts, and diagrams should be professionally drawn on plain white paper or tracing cloth in black India ink twice the size they are to be printed. A full page diagram, in print, measures 7.25 × 4.75 inches.

As far as possible, formulas should be typewritten and symbols not available on a typewriter carefully inserted in ink. Authors are asked to keep in mind the typographical difficulties of complicated mathematical formulae. The difference between capital and lower-case letters should be clearly shown; care should be taken to avoid confusion between such pairs as zero and the letter O, the numeral 1 and the letter l, numeral 1 used as superscript and prime ($'$), alpha and a, kappa and k, mu and u, nu and v, eta and n, etc. Subscripts or superscripts should be clearly below or above the line. Bars above groups of letters (e.g., $\overline{\log x}$) and underlined letters (e.g., $\underline{x}$) are difficult to print and should be avoided. Symbols are automatically italicized by the printer and should not be underlined on manuscripts. Boldface letters may be indicated by underlining with a wavy line on the manuscript; boldface subscripts and superscripts are not available. Complicated exponentials should be represented with the symbol exp particularly when appearing in the text, that is,

$$\exp\,[(a^2 + b^2)^{1/2}] \text{ should be used in place of } e^{(a^2+b^2)^{1/2}}.$$

In writing square roots the fractional exponent is preferable to the radical sign. Fractions in the body of the text (and when possible in displayed expressions) and fractions occurring in the numerators or denominators of fractions are preferably written with the solidus; thus

$$(a + b)/(c + d) \text{ rather than } \frac{a+b}{c+d}.$$

Authors will ordinarily receive only galley proofs. Fifty offprints without covers will be furnished free. Costs for additional reprints and covers can be furnished on request.

CONTENTS

TECHNOMETRICS, VOL. 2, NO. 2, MAY 1960

Rejection of Outliers ..... *F. J. Anscombe* 123

Locating Outliers in Factorial Experiments ..... *C. Daniel* 149

Discussion of the Papers of Messrs. Anscombe and Daniel
W. H. Kruskal, T. S. Ferguson, J. W. Tukey and E. J. Gumbel 157

Tests for the Validity of the Assumptions that the Underlying Distribution of Life is Exponential, Part II ..... *B. Epstein* 167

Partial Duplication of Response Surface Designs .... *O. Dykstra* 185

A Rank Sum Test for Comparing all Pairs of Treatments
R. G. D. Steel 197

The Percentile Points of Distributions Having Known Cumulants ..... *R. A. Fisher and E. A. Cornish* 209

An Approximation to the Negative Moments of the Positive Binomial Useful in Life Testing
W. Mendenhall and E. H. Lehman, Jr. 227

Order Statistics from the Gamma Distribution .... *S. S. Gupta* 243

Parallel Fractional Replicates ..... *C. Daniel* 263

Notices 269